Strategic Research and Innovation Agenda 2025

INTRODUCTION AND OVERVIEW

0. Introduction and Overview

0.1. PART 1 – WHY? GOALS AND PURPOSES OF ECS SRIA 2025

We are living in a period characterised by unprecedented challenges: climate crisis, ageing population, geopolitical tensions, to name a few. In particular, Electronic Components and Systems bear the promise to be key tools to allow us to overcome many of the aforementioned challenges. However, faced with a flurry of possible technology options, and limited financial and human resources, Europe needs to make sure that it aligns its research efforts and public support where it can best leverage its strengths and mitigate its weaknesses, both now and in the future.

This is the purpose of this ECS Strategic Research and Innovation Agenda (ECS SRIA), jointly developed by the experts of the European ECS community, coordinated by the three industry associations: AENEAS, Inside Industry Association (formerly ARTEMIS-IA) and EPOSS. This eighth edition describes the Major Challenges, and the necessary RD&I efforts to tackle them, in micro- and nanoelectronics for smart systems integration, all the way up to embedded and cyber-physical systems, and systems-of-systems (SoSs). By doing so, it is an essential tool to drive the ECS research and innovation efforts in Europe with the ultimate goal of creating value, growth, jobs and prosperity.

0.1.1 The ECS ecosystem: a key enabler of prosperity for Europe

The importance for Europe of the ECS ecosystem is twofold: On the one hand, it is a major contributor to European economic and political strength, due to its economical weight in terms of employment and wealth creation, and its contribution to European economical sovereignty. On the other hand, ECS bring benefits across many - if not all - aspects of our daily life.

A Strategic advantages for the EU

Globally, the long-term market trend for electronic components is expected to exceed US $1,000 billion by 2030[1]. In Europe, the semiconductor ecosystem employs some 250,000 people, with 2.5 million in the overall value chain of equipment, materials, semiconductors components, system integration, applications and services – mostly in jobs requiring a high level of education.

[1] Europe’s urgent need to invest in a leading-edge semiconductor ecosystem - Article - Kearney

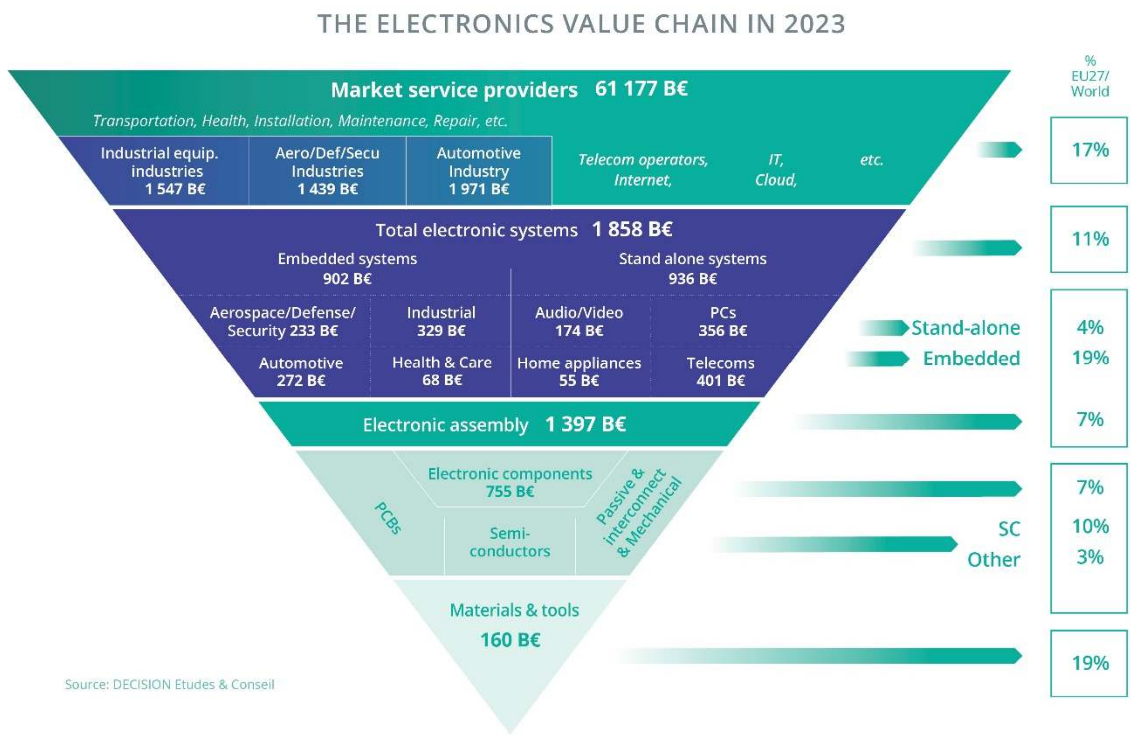


Figure 0.1 Process technology, equipment, materials and manufacturing is at the base of the digital value chain (Source: DECISION Etudes & Conseil) – 2023 market size numbers

The demand for chips is expected to double by 2030, driven by the digitalisation of society and the pervasiveness of Artificial Intelligence.

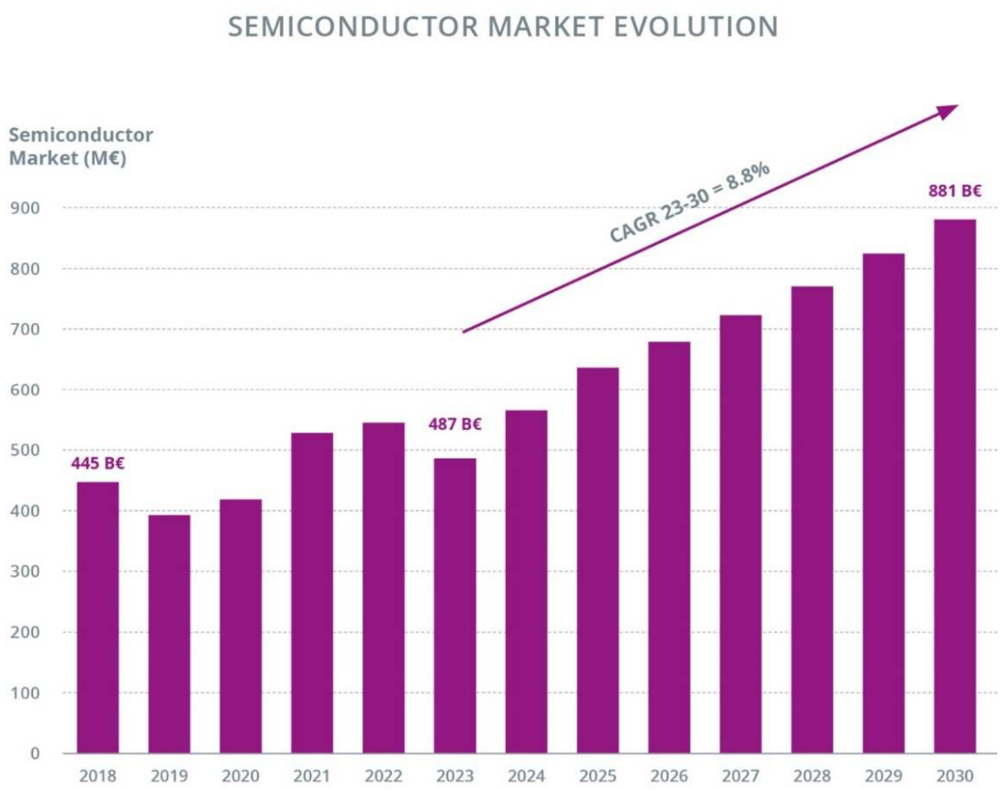


Figure 0.2 Semiconductor market evolution 2018 – 2030 (Source: DECISION Etudes & Conseil, WSTS)

Given that in the next decade, 85% of overall global growth is projected to take place outside the EU[2], it is essential that the EU maintains and increases the worldwide competitiveness of its electronics component industrial ecosystem. In turn, this increased competitiveness will stimulate all the ECS value chain and downstream industries that depend on it, including transportation, healthcare, energy, security, and telecommunications, to name a few. It is a key ingredient to maintain manufacturing as the backbone of the European economy, to guide further digitalisation of society and industry, ensure Europe's sovereignty and resilience against crises situations, facilitate the digital single market, maintain employment and move towards a better, smarter society.

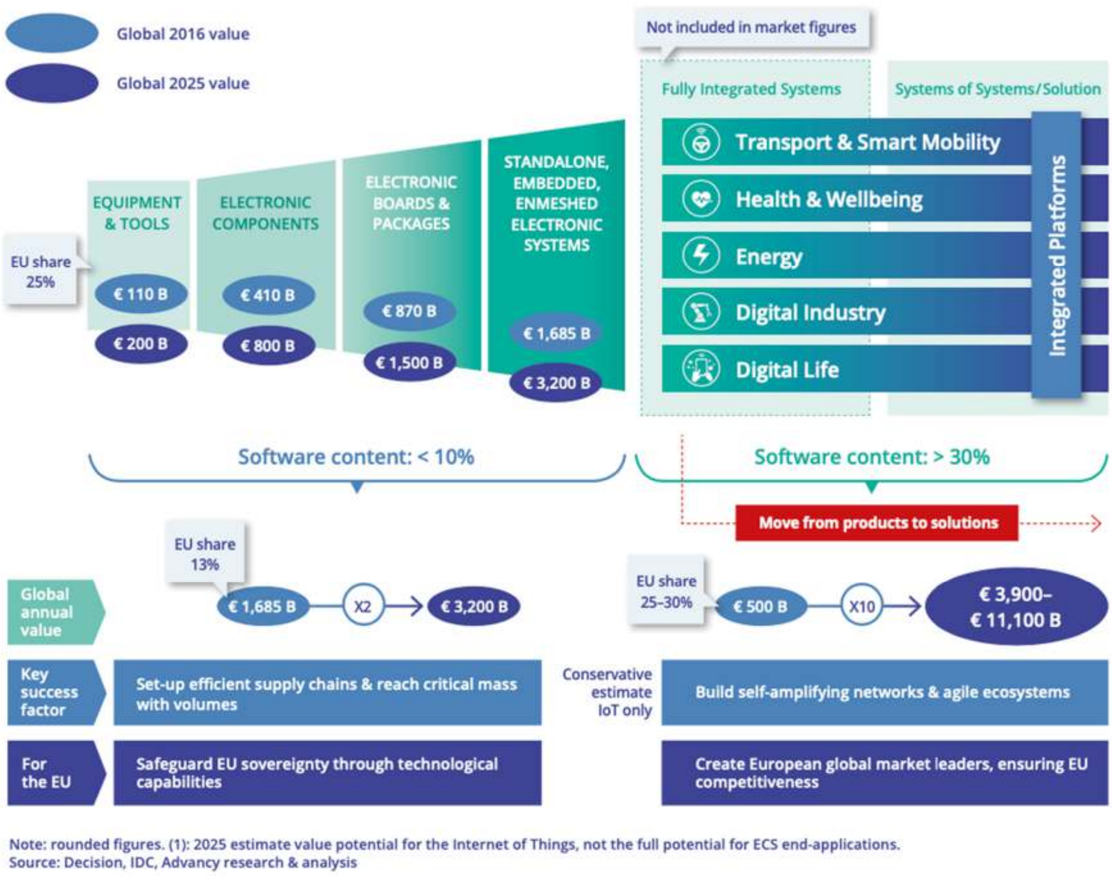


Figure 0.3 - Global and European value chain 2016–25[3].

[2] European Parliament: EVP Dombrovskis speech at hearing Commissioner-Designate for Trade https://ec.europa.eu/commission/commissioners/2019-2024/dombrovskis/announcements/european-parliament-evp-dombrovskis-speech-hearing-commissioner-designate-trade_en

[3] Source: Embedded Intelligence: Trends and Challenges, A Study by Advancy, Commissioned by ARTEMIS Industry Association, March 2019

Maintaining competitiveness of the European ECS value chain

Even though the European microchips manufacturing market share is only around 10 %, the European share of integrated products and electronic systems is much higher, for example in the automotive and telecom sectors. Many European companies master specialties technologies, including integration of those technologies into smart systems, combining hardware and software - including firmware and middleware -, which allow them to grab a relatively high share of the end-product value.

Systems-of-systems[4] and their formalisation were originally conceived and studied in the defence domain, but they are (and will be) vital infrastructure for many other vertical domains. For example, the shift in the mobility sector towards electrification and autonomous mobility necessitates the adoption of systems-of-systems in e.g. vehicles and roadside infrastructure. Given current and future expectations of the market, investment in SoS research and innovation[5] is essential to European leadership in the mobility sector. Likewise, to remain at the state-of-the-art in embedded systems architecture and software Europe should continue to invest in this domain, despite fierce competition. From this perspective, the convergence between AI and edge computing - embedded intelligence - should be a top priority.

Europe is also internationally known for its high-quality products. European engineers are highly skilled in systems engineering, including integration, validation and testing, thus ensuring system qualities such as safety, security, reliability, etc, for their products, using development and test tools and frameworks enabling them to ensure these qualities in an effort- and cost-efficient manner following international and European standards. The EU has a robust and reliable safety and product liability regulatory framework, and a rigorous body of safety standards, complemented by national, non-harmonised liability regulations. This ability to provide high quality products at affordable costs has led to trustworthiness with customers on the one hand and increased competitiveness on the other hand, which have been big success factors for European embedded systems in almost all industries.

Europe should take benefit of its specific strengths, and of its ambitious plans - such as the "European Green Deal" - to make its ECS industry even more sustainable and competitive producing trustworthy products.

Sovereignty

European strategic autonomy in ECS calls for further collaboration between European companies and organizations. Design frameworks, reference architectures and integration

[4] A collection of independent and distributed embedded and cyber-physical systems dynamically composed to generate a new and more complex system, provided with new functionalities and driven by new goals not present in the constituent embedded and cyber-physical systems individually. See full definition in the glossary

[5] From Internet of Things to System of Systems – Market analysis, achievements, positioning and future vision of the ECS community on IoT and SoS, P. Azzoni, Artemis 2020.

platforms will lead to new design ecosystems. Integration platforms[6] will provide the opportunity to leverage a high number of small and medium-sized enterprises (SMEs) and larger businesses into a platform-based economy mirroring the existing highly successful platforms of, for example, Google and Apple.

The above holds in particular for EDA Design Platforms, where global, non-European players like Synopsis and Cadence rule an overwhelming part of the market and thus de-facto control if, where and by whom such ecosystems can evolve. Providing European alternatives for such platforms will support technical enhancements e.g. the development of edge AI, embedded AI and embedded computing chips, support of the Open-Source Hardware Community (i.e., RISC-V), and many others. It will also facilitate the development of ecosystems, e.g. allowing to have a one-shop entry for start-ups/SMEs and academia to validate their innovations and new architectures into silicon, providing non-differentiating IP's, tool support, and a coherent design environment.

The Green Deal objectives will further drive the production of renewable energy and the large-scale adaptation of (bi-directional) battery electric vehicles and heat pumps. Consequently, Europe will have to modernize its energy grid towards a highly dynamic, blackout-protected energy infrastructure, also significantly reducing the dependency on imported fossil fuels.

Advancements in ECS, particularly in edge AI computing and in mastering the integration task into its products, will substantially contribute to enabling European Industries to build systems with guaranteed quality properties. Europe's strengths in creating trustworthy dependable systems by high-quality system design ("made in Europe" quality) contributes to European strategic digital autonomy.

Resilience

The Covid-19 pandemic has revealed the vulnerability of global, distributed value chains, with a disturbing and costly impact on society. In order to mitigate the impact of such disasters it is essential that strategic industrial ecosystems receive the backing from the EU. New models that will bring greater efficiency and more agile production processes need to be developed, and European manufacturing must be strengthened in key areas. This will ensure an effective and swift reaction to sudden market shocks as well as flexible manufacturing, accommodating shorter life cycles of products and fabrication-on-demand. Again, ECS innovations will play a key role here.

In semiconductor manufacturing specifically, Europe can reinforce its lead in semiconductor processing and packaging, equipment and smart systems based on the priorities set out in this ECS SRIA. The first Important Project of Common European Interest (IPCEI) on microelectronics, for example, was a successful step towards strengthening European semiconductor manufacturing in strategic areas where large-scale subsidies in other regions

[6] An integration platform is an ECS allowing the integration of different systems, applications and services into a single system. See full definition in the glossary

have started to threaten the position of European players. The European Commission has set ambitious targets[7] in its 'Digital Compass' to double the 'cutting-edge semiconductor' manufacturing share in Europe in 2030 – to maintain strategic autonomy, and to be involved in AI and other key technologies of the digital world.

Further, the EU can help national governments, companies and citizens to cooperate more easily, and develop reliable societal emergency infrastructures. This will make European societies better prepared to deal with emergency and crisis situations.

Employment challenges

The European electronics industry is currently facing a skill gap. Sufficient numbers of engineers with the right skills are crucial for Europe to compete with other regions and exploit the sector's true potential for the European economy. Faced with these challenges, Europe can play on two levers to develop a strategic advantage vs. other regions of the world:

- On the one hand, it must maintain and strengthen its traditionally strong and advanced educational system, and the presence of world-leading research institutes throughout the whole stack of competencies. Universities have a vital role in the supply of graduate engineers, and it is essential that graduates have access to industry-relevant design tools, leading edge technologies and training. Programs such as EUROPRACTICE are essential in providing this access in an affordable manner. Likewise, continued investment in semiconductor-related studies, as intended under the Pact for Skills[8], is crucial to reversing the current trend of declining numbers of students.
- On the other hand, there is a need to drastically increase the efficiency of engineering activities throughout the whole design / develop / test / validation process by providing new tools that are able to handle new product features which in turn are enabled by usage of AI and other new technologies. Model-based and AI-supported technologies will contribute considerably to the increase in efficiency, the increase in capabilities, the mitigation of shortage in software engineering resources, and further improve overall software consistency and quality.

While ECS is facing a skill gap, other industries are at risk of suffering job losses. For example, Europe's automotive industry employs around twelve million people. Currently, the transportation sector is undergoing a fundamental and complex transformation across all modes. Its position is challenged by US information technology giants and aggressive and new agile Chinese automotive companies. The ECS community will contribute substantially to maintaining competitiveness, and therefore jobs, of the European car industry by using new technologies, components and systems to target areas like autonomous vehicles, electrified CO_2-neutral vehicles, Over-the-air (OTA) updates and new mobility concepts to reduce overall energy usage (e.g. mobility as a service).

---

[7] Europe's Digital Decade: digital targets for 2030 | European Commission (europa.eu)

[8] Pact for Skills - Employment, Social Affairs & Inclusion - European Commission (europa.eu) Pact for Skills - Employment, Social Affairs & Inclusion - European Commission (europa.eu)

Conclusion

The global digitalisation of society and industries is enabled by ongoing innovations in ECS. With its specific strengths Europe is well positioned to prevail in the fierce global competition. We are strong in embedded systems architecture and software, in embedded intelligence, in specialties semiconductors (Power, RF, sensors, FD-SOI, MEMS…), and in maintaining overall high quality, safety and security standards in ECS. Therefore, the European ECS market prospects are seen as strong.

Further investments in the future of electronic components, modules and systems integration have the following strategic advantages for Europe:

- Strengthening Europe's economy through the generation of high-tech innovations.
- Increasing the added value in Europe by integrating more functional systems and products, e.g. in automotive, med tech and telecom industries.
- Enabling a successful Twin Transition (Green and Digital) in an economically feasible way through multifunctional smart devices.
- Ensuring European sovereignty and securing strategic Intellectual Property from European companies on advanced technologies in the microelectronic ecosystem with regards to heterogeneous integration.

To maintain leading positions in specific markets Europe must continue to invest in RD&I in strategic ECS technologies, including disruptive technologies like high NA lithography, 2D-materials, quantum computing and related technologies, AI based development, design and verification methodologies, scalable platform technologies and differentiating capabilities by software. With strategic initiatives like the Green Deal, the Digital single market, the Pact for Skills, and the Chips Act, to name a few, the European Commission has already taken important steps to secure a healthy future for the ECS industry in Europe, and therefore for many ECS-enabled end-product industries. The ambition of the ECS Strategic Research and Innovation Agenda is to support, enable and amplify that effort, by:

- Strengthening industry involvement: Strategic R&D and manufacturing investments in disruptive technologies, needed to cope with the transition challenges of our society (digitalization and sustainability, EU sovereignty), as depicted in the ECS SRIA, need collaborative approaches for R&D and governmental support for lowering the risks for realization and implementation. Therefore, in particular industries with high R&D intensities addressing key enabling technologies to solve global challenges, such as Electronic Components and Systems (ECS), should be supported to grow;
- Providing guidance to join forces: For our continent with federal democratic structures and a high degree of individual required solutions, joining forces in a billion-market environment is the only viable approach to be well set-up for the future despite having limited resources. Collaborative RD&I projects are one of the main instruments to that effect in our federal eco-system.

B Technology-enabled societal benefits

Besides its economical weight per se, the strategic importance for Europe of the ECS ecosystem is further strengthened due to its role as the enabler of the digitalisation of our society.

Digitised services based on smart electronic components and systems (ECS) are becoming increasingly ubiquitous, penetrating every aspect of our lives. They also provide ever greater functionality, more connectivity and more autonomy. The benefits to society are numerous: safer traffic, less carbon emission and pollution, instantaneous access to information, convenience and cost saving in e-health, more efficient factories, more social safety, and many more. In short, ECS-based applications and services are key to ensuring the stable growth and development of the European Union.

On the other hand, our dependency on ECS-based applications is also continuously growing. This reinforces people's concerns on – among others – the security of privacy and personal possessions (e.g. through cybercrime, cloud services, surveillance cameras), on personal safety (driver-assistance, e-health), and on the unclear impact of transformational technologies (AI, quantum technology). The societal concerns arising from these technologies need to be addressed. To achieve that, it is vitally important that these electronic systems are safe, secure and trustworthy. A human-centred approach is a key aspect of the EU's policy on technology development, and in line with fundamental European social and ethical values.
In the following pages, we examine in more details the societal benefits which ECS brings to Europe, and where R&I efforts are needed to ensure further gains.

Sustainability

Contribution of ECS technologies to sustainability are two-fold:

On the one hand, ECS are a major enabler of the Green Deal objectives as they foster sustainable and smart mobility; the supply of clean, affordable and secure energy, resource-efficient manufacturing, a healthy and environment-friendly food supply chain, healthcare transformation, as well as more efficient information systems.

To accelerate the shift towards sustainable and smart mobility, the mobility sector requires energy efficient mobility grade semiconductor devices, be it for vehicles or for the supporting infrastructure required for the charging of EV's. This needs to be complemented with middleware and embedded SW applications, which are often part of a cloud-to-edge continuum.

In the energy sector, the power grid architecture will be transformed into a multi-modal energy system architecture. It will comprise distributed renewable energy generation, energy conversion units for sector coupling, transmission and distribution grids allowing bi-directional power flow, and storage for all modes of energy (electric, thermal, chemical). Key to these new energy applications will be smart sensors, networks of sensors, and smart actuators that enable status monitoring on each grid level, as well as smart converters (for all voltage levels) based on highly efficient and fast-switching semiconductor power devices and modules that enable real-time control of energy system components and grids for optimised operation based on forecasts of generation and demand, but also in case of any critical event. The future grid operation, finally, requires a sophisticated information and communication infrastructure including cloud services, IT security, and AI technologies.

Besides, ECS will be at the core of implementing new ways of manufacturing, to promote an environmentally friendly production and to develop building blocks in advanced automation and control, advanced sensors, digital twins, artificial intelligence, collaborative robotics, monitoring through value chains – that will allow for better accuracy and performance, better (predictive) maintenance and higher asset utilization. It is particularly important to ensure that design processes will cover the complete lifecycle of products for future ECS-based applications. Data must be collected to this aim, and used to enable continuous updates and upgrades of products, but also in-the-field tests of properties that cannot be assessed at design-, development- or testing-time. This will as well increase the effectiveness of validation and test steps by virtual validation methods based on this data.

Another major application sector for ECS is agriculture, where policies focus on preservation of landscapes, biodiversity, and environmental protection. This will require measurement and monitoring technology which is accurate, highly scalable, and secure. In addition to these environmental requirements, Smart Internet of Things (IoT) systems support productivity growth, access to clean water, fertile soil and healthy air for all, and help fighting against pests while preserving biodiversity and restoring the planet’s ecosystems.

Healthcare digitization will enable the shift from hospital care to remote care at home, thus reducing travel-induced environmental impact, while personalized medicine will reduce waste in resources.

Finally, edge computing and embedded intelligence will allow to significantly reduce the energy consumption for data transmissions (e.g. to the cloud), will save resources in key domains of Europe’s industrial systems, and will improve the efficient use of natural resources.

On the other hand, the ECS sector is focusing on improving its own energy performance and disposability of electronic components and reducing its environmental footprint by means of cleaner and greener production processes, more circularity and less energy and material consumption. These efforts are mandatory since the demand for ECS is set to grow considerably to serve environmentally sustainable applications, and we cannot allow this to translate into an equivalent increase in the resources consumed in their production and operation.

The sustainable manufacturing of semiconductors requires the continuation of significant R&I on new processes, manufacturing techniques, equipment and materials. Advances in the manufacturing of chips and packages in the coming years will strongly contribute to Europe’s ambition to become climate-neutral by 2050. Some examples are:

- Device scaling - by moving into 3D for sub-3 nm node memory - and computing technologies will drive down energy consumption following the power, performance, area and cost (PPAC) scaling roadmaps.
- New embedded non-volatile memory technologies enable local processing and storage of configuration data, decreasing data transmission and energy needs for a wide range of automotive and IoT applications.
- New power electronics devices, either based on silicon or new (GaN, SiC) materials, will increase the energy efficiency of electric powertrains, energy storage, lighting systems, etc.
- Improved integration technologies and miniaturisation will support sustainability of products and production technologies.

One key point of attention is the development and use of replacement materials to comply with Restriction of Hazardous Substances Directive (ROHS) regulations (such as lead, mercury and other metals, flame retardants and certain phthalates, PFAS[9]) and minimization of critical raw materials (CRM) dependence. In particular, due to their unique physicochemical properties, PFAS are currently extensively used in both chip fabrication and for semiconductor manufacturing equipment and factory infrastructure. Research to identify alternative chemistries and to develop efficient abatement technologies for uses where no alternative can be found is therefore essential.

Beyond the sustainability of the components, modules and systems fabrication stage, the full life cycle and end-of-life of ECS should be considered using life-cycle assessment (LCA) as a design tool. There must be upstream considerations and design for sustainable production and condition monitoring, HW and SW upgradeability, lifetime extension, health monitoring, predictive maintenance and repair. These are all to be integrated into a lifecycle-spanning continuous development process that enable feedback of data from the later phases of the lifecycle to the former ones. Besides, reduction of CO_2 emissions during the lifetime of the system requires minimizing the power consumption at component, module and system levels while in operation by using low-power hardware and software technologies.

Moreover, eco-design has to consider re-use in a second life application, and the recovery of components and materials for recycling. Note that increased integration will cause the borders between components, modules and systems to become blurred: More diverse and complex materials will be used at each level, so that the dismantling of systems into their constituent components at the end of their useful life will become increasingly difficult, requiring significant R&I efforts to identify solutions to achieve circular economy.

Finally, another lever to decrease the overall resource consumption of ECS is at the architectural level. Embedded intelligence for instance can provide computing capabilities to the nodes and devices of the edge of the network (or edge domain) to improve the performance (energy efficiency, latency, etc.), operating cost, and reliability of applications

---

[9] PFAS is a class of thousands of synthetic substances known as 'forever chemicals' since they do not break down in the environment. PFAS offer a unique set of technical characteristics, which include exceptional heat and chemical resistance, high electrical insulation resistance, high purity, low-outgassing and low coefficient of friction

and services. Above all, this new computing paradigm could significantly reduce environmental footprint by the introduction of ultra-low power and efficient computing solutions. There is plenty to explore on the trade-off between performances and power consumption reduction, and on managing complexity (including security, safety, and privacy) for embedded architectures to be used in different applications areas, which will spread the use of edge computing and artificial intelligence and their contribution to European sustainability.

Healthier life

The critical role that smart components, modules and systems can play for the world's security and health was demonstrated during the Covid-19 pandemic. Key topics here range from an acceleration in the analysis of DNA samples, the availability of automated medical support and diagnosis tools, to tracking systems for tracing and controlling the spread of the disease, as well as the recourse to robots in several hazardous situations, from disinfecting airplanes and hospital rooms, to delivering medication to isolated patients.

Even outside times of crises, advances in ECS are playing a key role towards enabling a healthier life for all citizens. The Internet of Things (IoT) is one of the main technologies enabled by smart components, modules and systems, and ongoing advancements in IoT bring more smartness to people's health and well-being via e-health, m-health, implants, ingestibles, wearables, personalised medicine, inclusion of people with a disability, etc. This trend is further accelerated by many AI-enabled positive breakthroughs which can be seen on the horizon.

The potential of these technologies is enhanced by their integration into complex systems-of-systems, relying on ubiquitous, wideband and dependable connectivity, improving medical practices and services for patients and healthcare professionals. Connected devices can allow remote monitoring and diagnosis and more efficient means of treatment. Access to healthcare for rural populations will also improve.

Smarter, more efficient society

AI and edge computing have become core technologies for the digital transformation. AI will allow to analyse data on the level of cognitive reasoning to take decisions locally at the edge (embedded AI), transforming the IoT into the Artificial Intelligence of Things (AIoT). Likewise, control and automation tasks, which are traditionally carried out on centralised computer platforms, will be shifted to distributed computing devices, making use of decentralised control algorithms.

Embedded cyber-physical systems (ECPS), or IoT systems, provide data processing and intelligence on the site/edge, while improving security and privacy, and completely changing the way we manage everyday activities. ECPS also play a critical role in modern digitalisation solutions, quickly becoming nodes in distributed infrastructures supporting systems-of-systems (SoSs) for monitoring, controlling and orchestrating supply chains, manufacturing lines, organisation's internal processes, marketing and sales, and consumer products.

Moreover, digitalisation platforms exploit embedded software flexibility and ECPS features to automate their remote management and control through continuous engineering across their entire lifecycle (e.g. provisioning, bug identification, firmware and software updates, and configuration management).

Ongoing advancements in IoT systems will drive the further digitalisation of society, by bringing more smartness to human activities, like smart cities, smart transportation, smart grids, smart manufacturing, and pull the development of cyber physical systems and embedded systems-of-systems (SoS). A clear example is the introduction of autonomous vehicles, which will become components in the complex logistics systems of cities, countries, and regions. SoS-related technologies will be key to providing efficient utilisation of autonomous vehicle assets, also offering timely delivery of goods and personnel. Another example is the integration infrastructures adopted in production, in order to meet customer demands locally. Here, the interoperability of SoS technologies across domains is an essential capability.

This evolution will in turn put high demands on the availability and reliability of high-speed, secure, low or guaranteed latency connectivity.

Another important trend is the emergence of open-source components which become the core building blocks of application software in many innovative domains[10]. Developers are being provided with an ever-growing selection of off-the-shelf possibilities, which they can use for assembling their products faster and more efficiently, where efficiency goes hand in hand with affordability and sustainability.

Privacy, Safety and Security

The efficiency and flexibility of embedded software, in conjunction with the hardware capabilities of ECS, allow to move various processing functions to local devices, such as voice and environment recognition, allowing for privacy preserving functionalities. AI also increase the capabilities to detect intrusions, thus reinforcing the protection of privacy.

Furthermore, the ever-improving detection capabilities of cameras and other long-range sensors (radar, lidar, etc…) combined with the development of AI at the edge is opening unprecedented opportunities for many safety and security-related applications that currently rely on human involvement, such as automated driving, security and surveillance and process monitoring.

Connected functionalities will extend the control and automation of a single system (e.g. a truck) to a network of systems (e.g. a truck platoon), resulting in networked control of a cyber-physical system. The benefit of this is generally better performance and safety.

As already noted regarding the support for a smarter, more efficient society, the technical trends underlying the promises for higher privacy, safety and security demand the availability

[10] More than 81% of produced software are consuming open-source code in products or services (https://github.com/todogroup/osposurvey/tree/main/2020)

of ever better-performing connectivity networks. As an example, automated driving requires ultra-high reliability, extremely low latency and high throughput connectivity solutions. Advanced edge solutions that will integrate AI/ML schemes over secure links will also be of paramount importance. Advanced connectivity is also a key enabler for disaster relief and prevention systems.

Finally, one cannot ignore that the ever-increasing importance of ECS-based systems in our daily lives raise concerns regarding trustworthiness, privacy, safety and cybersecurity. Indeed, a degraded behaviour of cyber-physical systems or an incorrect integration among them, would affect vital properties and could cause serious damage. Shortcomings in those dimensions might even outweigh the societal and individual benefits perceived by users, thus lowering trust in, and acceptance of, new technologies. Ensuring high standards on safety, security, and reliability at affordable efforts and cost, require continuous R&I efforts on methods and tools for ECS architecture and design:

- Applying "quality by design" approaches for future ECS-based systems.
- Providing methodology, modelling and tool support to ensure end-to-end trustworthiness of new designs. This includes balancing trade-offs of quality aspects and ensuring tool-supported verification and validation (V&V) at the ECS level, providing methodology, modelling and tool support to validate safety of AI-based systems.

Connected society / social inclusion

The further integration of "smart everything" into "ubiquitous smart environments" will introduce large and very complex systems of systems with complex physical interactions. In this context the technology competence and innovation in the field of embedded and cyber-physical based SoS will be a critical asset.

ICT technologies have long been recognised as promoting and facilitating social inclusion, as well as digital inclusion (i.e. the ability of people to use technology). These aspects span dimensions as diverse as disaster relief, food security and the environment, as well as citizenship, community cohesion, self-expression and equality. One illustration is the use of AI to enhance conversational interfaces, improving the human-machine interface with reliable understanding of natural language.

0.1.2 The ECS SRIA: An agenda aligned with European priorities

For each technology and application domain it covers, the ECS SRIA identifies specific so-called Major Challenges, with a focus on the most critical aspects to be tackled from the perspective of innovation. The analysis of each Major Challenge illustrates the state of the art of the associated technology and/or application domains, describes the vision of the ECS community for the future, identifies potential outcomes, and defines research and engineering activities on the key focus areas that are fundamental to successfully address the challenge.

ECS are the key technology, enabler and differentiating factor behind all mass applications and their innovations, serving a market of over 61 trillion Euro (see figure 0.1). Therefore, a very broad scope in the applications, in the design, the underlaying technologies, the requirements and in the manufacturing is addressed in this document, resulting in the identification of 71 Major Challenges.

They are frequently interdependent – they influence each other, become increasingly demanding, and have an impact on many areas, including technology innovation, industrial competitiveness, security, safety, business and environmental sustainability, society, etc. From this perspective, the Major Challenges represent key factors for the achievement of four Main Common Objectives, which are aligned with the European Commission's strategic priorities (see table in Appendix):

- Boost industrial competitiveness
- Ensure EU digital autonomy
- Establish and strengthen sustainable and resilient value chains
- Unleash the full potential of intelligent and autonomous ECS-based systems

A Main Common Objective 1: boost industrial competitiveness through interdisciplinary technology innovations

As mentioned earlier, investing in the Research and Innovation topics described in this document will not only translate into a more competitive and sovereign European ECS value chain, but it will also boost European industrial competitiveness across all sectors.

Electronic components and systems, by their inherent nature, are the result of interdisciplinary research and engineering. These require competencies in diverse technology domains, including process technology, equipment, materials and manufacturing, electronics, and telecommunications, as well as cross-sectional technologies such as edge computing, artificial intelligence, high-speed connectivity, and cybersecurity.

As a result, ECS research needs to be interdisciplinary to benefit from the multiple available sources of innovation, as well as research-intensive and market-oriented. This will ensure forthcoming ECS innovations will be of strategic value for Europe and boost its industrial competitiveness in all its value chains, and help building the strong industrial base essential for European strategic autonomy.

With the Chips Act, the KDT JU has transited to the Chips JU, with a mandate extended to capacity building activities and related research and innovation activities of four operational objectives of the Chips for Europe Initiative, as set out in the Chips Act article 4:

1. building up advanced design capacities for integrated semiconductor technologies;
2. enhancing existing and developing new advanced pilot lines across the Union to enable development and deployment of cutting-edge semiconductor technologies and next-generation semiconductor technologies;
3. building advanced technology and engineering capacities for accelerating the innovative development of cutting-edge quantum chips and associated semiconductor technologies;

4. establishing a network of competence centres across the Union by enhancing existing or creating new facilities.

These objectives translated in particular into the following additional activities for the JU:

- The development of new capabilities to design, prototype and test innovative chips through pilot lines on semiconductor technologies, with the close collaboration of European RTOs and industrial vertical sectors.
- The establishment of competence centres, which will address the lack of skills that is affecting the European labour market, trying to provide training support for future professionals in the whole ECS domain, consolidating and strengthening the rich knowledge base required by interdisciplinarity.
- The setting up of a new Virtual Design Platform and specific funding instruments which will provide support to the rich, diverse and multidisciplinary ecosystem of European SMEs and start-ups, facilitating investments in research, development, and production and simplifying the process of bringing innovation from "the lab to the fab".

This SRIA is the reference document for all research and innovation activities of both the former (under KDT) and these new Chips JU activities. This extension reflects the intrinsic interdisciplinarity of the SRIA and further strengthens it from different perspectives.

All the new initiatives introduced by the Chips Act have been conceived to boost European competitivity in the ECS market and ECS-based application domains, guiding investments and resources towards strategic areas of semiconductor technology that are most critical for the European industry's competitiveness. This SRIA contributes to the identification of these critical areas and of the associated challenges that will characterise their development in the next ten years. The role of the SRIA is fundamental to identify the key starting points for RD&I and to setup multi-annual synergies with the pilot lines that will focus on more mature prototyping. The SRIA will also potentially provide the elements for the "first-of-a-kind facilities" envisioned by Pillar 2: the synergies between these three steps represent the key strategy to increase European capacity building, to boost competitiveness, and to anticipate, prevent and effectively manage future crises.

B Main Common Objective 2: ensure EU digital autonomy through secure, safe and reliable ECS supporting key European application domains

The benefits for European strategic autonomy of the development of innovative ECS technologies focused on security, safety, reliability, dependability and privacy was discussed in the section "strategic advantages for the EU". European technology-based, secure, safe and reliable ECS, combined with European AI solutions, are critical to securing global leadership and strategic autonomy in key areas such as ICT and to ensure compatibility with EU values.

These innovative technologies will simplify the implementation of the European Strategy for Data[11,12], and ensure security, privacy-by-design and strategic autonomy all along the industrial and digital value chains.
Threats to Europe’s strategic autonomy are to be found in the microelectronics value chain, and then downstream in the component user segments of the electronics industry. In this context, the Major Challenges identified by the ECS SRIA will help develop innovations in secure, safe and reliable ECS technologies for creating EU-based/-made solutions in the key European application domains of:

- Aerospace, defence, security.
- Automotive, transportation.
- Machinery, robotics, electrical equipment, energy.
- Communications, computing.
- Healthcare and well-being, etc.

The Chips Act aims at consolidating and strengthening European strategic autonomy, extending and boosting the KDT with a new specific focus on the upstream of the ECS value chain. Reducing the dependency on non-European suppliers and avoiding future shortages represent the main steps towards strategic autonomy: reaching the independence on non-European suppliers for critical ECS is crucial to ensure European industries and key application domains to have a stable supply of semiconductors, even during future global disruptions or geopolitical tensions. Technological leadership and strategic autonomy are the primary ingredients required to:

- Control ECS manufacturing to produce solutions designed to meet stringent security requirements; an essential aspect for maintaining secure digital infrastructures and key applications (e.g. automotive, defence, energy, etc.).
- Ensure resilient supply of semiconductors produced in Europe to support the resilience of critical infrastructure, including energy, transportation, and healthcare systems, which rely on advanced chips.
- Achieve a greater control over the technology infrastructure that underpins the European digital economy, ensuring that critical data and systems remain under European jurisdiction.
- Enhance Europe's economic competitiveness, leading to the creation of high-tech jobs, stimulating innovation, fostering economic growth, and making Europe less dependent on external economic forces.

C Main Common Objective 3: establish and strengthen sustainable and resilient ECS value chains supporting the Green Deal

European strategic autonomy will also require the sustainability and resilience of the entire ECS value chain since the development of innovative technologies focused on sustainability and the Green Deal will support ambitions to achieve a green, resilient and competitive Europe.

---

[11] https://ec.europa.eu/digital-single-market/en/policies/building-european-data-economy
[12] https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:52021DC0118

Moreover, the serious effects of climate change that we are experiencing daily and the current geo-political situation, which highlights our dependency on non-European fossil energy providers, further reinforce the need for Europe to accelerate its transition to climate neutrality by 2050.

This challenge must be perceived as an opportunity to create a new environment for boosting innovative aspects of technology and business models through achieving the following:

- Relying extensively on ECS-based technologies and digitalisation as key factors for lowering our global energy footprint at all levels of the economy, and by placing sustainability at the heart of combined digital and green transitions.
- Positioning the European players in hardware as front-runners in sustainability to secure a wider market so they can become world leaders. This will need European companies to consider the circular economy, new market positioning (by turning small market shares into specialisation areas), the environmental impact of global manufacturing, etc.
- Establishing this carbon-neutrality challenge, based on a close link between the digital and green transitions at the core of future funded collaborative research and innovation in ECS. This will help ensure a positive impact for each stage of the value chain, and achieve carbon neutrality right down to the final application/digital service.

This new context is required to fight and reduce the effects of climate change.

As developed earlier in the section “Sustainability” of “Technology-enabled societal benefits”, advances in ECS technologies are both a major enabler of the Green Deal for all ECS-based application fields, and a direct contributor through their significant impact on power and resource consumption of ECS manufacturing and use.

Furthermore, the **strategic autonomy** introduced by the Chips Act contributes to **economic sustainability and security**, through a reduced dependency on non-European suppliers, an increased production capacity, more competitive manufacturing processes and products, a resilient supply chain, circular economy promotion, etc.

D Main Common Objective 4: unleash the full potential of intelligent and autonomous ECS-based systems for the European digital era

ECS must have intelligence and autonomy capabilities to control their complexity more efficiently and more cost-effectively. This will help provide novel advanced functionalities and services, limit human presence to only where it is strictly required, improve the efficiency of vertical applications, etc. Intelligence and autonomy are also required for the role of ECS in the application domains, representing an important factor for the sustainability and resilience of the value chains: an ECS-based system that provides intelligent energy management, relying on technologies such as AI, represents a key building block – for example, for smart home and energy applications. Moreover, it also improves the resilience required to ensure optimal energy consumption in critical conditions and contributes to the sustainability of the value chain associated with vertical applications, since it reduces operational costs and environmental impact, improves the quality of service (QoS), return on investment (ROI), etc., thereby strengthening the global competitiveness of European companies and helping to achieve the objectives of the EU’s Green Deal.

With an innovative and strong semiconductor industry Europe can improve its competitiveness in the global AI landscape, and this will contribute attracting AI talent, companies, and research initiative in Europe. Supporting the key focus areas identified in the SRIA will be crucial to consolidate the European future position in the AI market and contribute to the supply chain resilience, supported by a robust semiconductor industry in AI hardware. Boosting the RD&I activities described in Chapter 2.1 of this SRIA will ensure a solid support for all the vertical domains covered by the Application chapters.

0.2. PART 2 – WHAT ? CONTENTS OF ECS SRIA 2025

0.2.1 Scope

This SRIA is intended to be funding-programme-agnostic, and can be used as a basis for the various cooperative programmes across Europe.

However, the scope of our work, and of this document, is firmly within the ECS domain. For details on developments in the specific application areas further up the value chain, please consult the SRIAs of other associations or public/private partnerships (PPPs) addressing those specific areas.

The range of this ECS SRIA is very wide, going from transistors within silicon chips acting as individual electrical switches for integration in smart systems up to global system-of-systems performing complex cognitive tasks and interacting with numerous humans and machines over a wide geographical spread. A very simplified view of this ECS technology "stack" is illustrated in Figure 0.4 with an example.

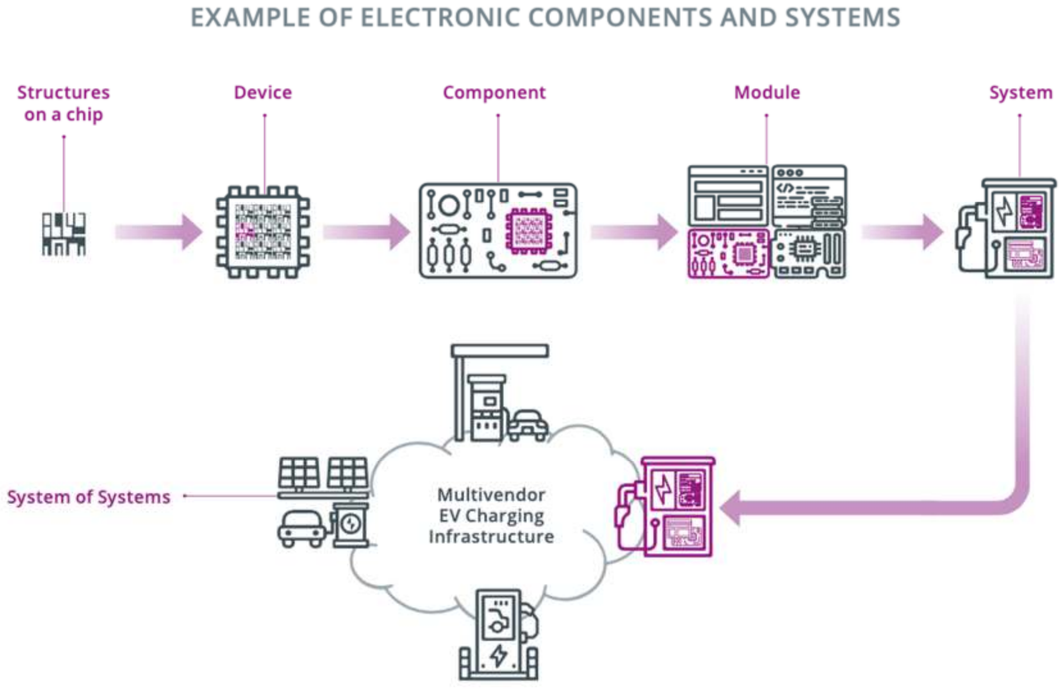


Figure 0.4 Different integration levels illustrated by the example of an EV charging infrastructure[13] (Source: Eurotech)

Designing such artefacts requires an interdisciplinary hierarchical approach, whereby various ECS specialists are working at different abstraction levels. As a result, the same term can have different meanings for specialists of different ECS domains: for instance, a

[13] Structure on a chip: elementary building blocks of an integrated circuit, such as a FDSOI or FinFET transistor, or more complex structures such as an embedded memory block.

"system" designed and implemented within a given development process may be integrated as a "component" into a higher-level "system" within another development step of the engineering process. Nevertheless, to avoid confusion, since 2021 the ECS SRIA includes a glossary, where many of the key terms are defined, to avoid inconsistency across the various chapters. It was also felt that developing a common language was important in building a strong and integrated ECS community. In addition, some of the bricks of the ECS technology "stack" are further detailed below.

- Device: in the context of the ECS SRIA, and if it is not further qualified, a device will be defined as a "packaged chip", whether it is a packaged integrated circuit (e.g. system on a chip, memory, processor, or microcontroller) or a micro-electromechanical system (MEMS)/micro-opto-electro-mechanical system (MOEMS). A device performs a general electrical, electronic or electrical/electronic/physical transduction role.

- Component: a combination of devices and other elements (such as passives) that fulfil a specific need, such as transduction of a single physical parameter within a well-specified case. A component is not self-contained in all its functions, as it requires the close support of other components for operation (e.g. in data processing, power handling, and embedded software).

- Module: a combination of correctly integrated components in which their assembly embodies a specific functionality required for the proper working of a system (e.g. sensing and actuation module, control module, communication module, energy provision module). A module is self-contained in hardware and software, making it interchangeable between systems, and allowing a higher abstraction level in systems design.

- System: for the purpose of this SRIA, a system is a set of electronic-based constituents (subsystems, modules and components, realised in hardware, software, or both) that are integrated in a way that allows the system to perform a desired (set of) function(s). Due to ECS typically being constructed hierarchically, a "module" (e.g. camera or other sensor) being part of the electronic "system" in an autonomous car might itself be referred to as a "system" when being designed (e.g. while integrating lower-level components together to achieve the "camera" function).

- System-of-systems (SoS): a collection of independent and distributed embedded and cyber-physical systems dynamically composed to generate a new and more complex system, provided with new functionalities and driven by new goals not present in the constituent embedded and cyber-physical systems individually. The difference between a "system" (comprising subsystems, modules and components) and a "system-of-systems" (also comprising subsystems) is that the subsystems of a system are chosen and integrated during design-time (i.e. completely under the control of the engineers), while in a system-of-systems the constituent (sub)systems are physically independent and dynamically form a system-of-systems at run-time.

0.2.2 The structure

The first part of the ECS SRIA is composed of four chapters focused on the **Foundational Technology Layers** and their technical challenges along the technology stack, from materials and process technology to components, modules and their integration into electronic systems, embedded software developments and software technologies, to full systems and systems-of-systems. These foundational layers are characterised by hierarchical dependencies due to the inherent nature of ECS and the way they compose and integrate in complex structures. Advances in all **Foundational Technology Layers** will be essential to creating new electronic chips, components, modules, systems, and systems-of-systems along the value chain: these are the fundamental elements required to build the digitalisation solutions of the future.

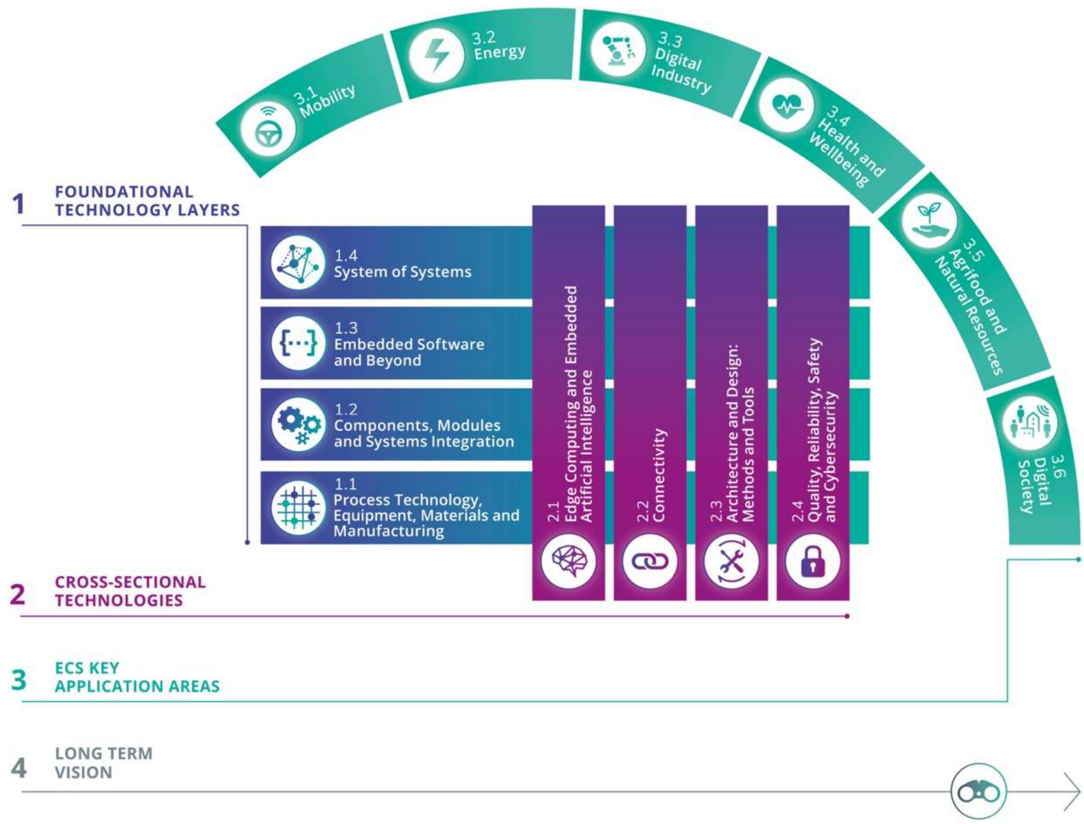


Figure 0.5 The structure of the ECS SRIA

The foundational layers represent a very fertile ground where new interdisciplinary technologies, products and solutions can grow. They are then complemented in the second part of the ECS SRIA by four **Cross-Sectional Technology** chapters that focus on transversal areas of scientific research and engineering, where innovative results emerge from the joint contribution of the foundational layers to those specific areas. **Edge Computing and Embedded Artificial Intelligence**, or **Connectivity** (e.g. 5G to 6G) will require new integrated circuits to develop innovative electronic components that can be used to develop smarter and more connected components, modules and entire systems, running smart software that will offer new functionalities and capabilities. That will allow these systems to interact, cooperate and merge in larger systems-of-systems. Similarly, **Architectures and Design: Methods and Tools** have to be further developed to provide support to each of the foundational layers, covering all domains along the technology stack, across the entire lifecycle of technologies and products. The same applies to **Quality, Reliability, Safety and Cybersecurity** concepts that can only be addressed successfully if they are encompassing the whole ECS process flow along the entire value chain.

The innovation generated by these cross-sectional technologies will be applied across foundational layer stacks and amplify the effect of innovation in all key ECS application domains. Of course, there is some overlap among the eight technology chapters since they are closely linked, but as they examine the individual challenges from different perspectives, this overlap is extremely constructive and generates valuable synergies.

In the third part of the ECS SRIA, six **Application** chapters describe the challenges arising from specific ECS application domains that are key for Europe and identify the RD&I efforts required by these application domains as regards ECS.

Finally, the **Long-Term Vision** chapter illustrates our vision of the ECS beyond the time horizon covered by the other chapters. It seeks to identify the research subjects that must be addressed at low TRL levels as foundation and preparation for the crucial developments in European industry over the next decade. Based on the trends and plans described in the preceding chapters, the long-term industrial requirements are also examined to help research programmes understand which hardware, software and system solutions should be produced most effectively for the continuous improvement of European digital technology.

While the overall structure of the ECS SRIA is unchanged with respect to the previous edition, the internal chapter structure has undergone a significant revision: In particular, the sections "Strategic advantages for the EU" and "Technology-enabled societal benefits", which were scattered across chapters in the previous edition, have been regrouped and summarised, to be now part of this introductory chapter. The rationale behind that restructuring is that many arguments developed in those sections were not chapter-specific but rather generic in nature. This allowed to avoid redundancies, refocus the messages of the individual chapters on their specificities, and reduce the overall ECS SRIA text size.

In the same spirit, the chapter editing teams have been reviewing their text with the goal to achieve greater conciseness, while of course eliminating outdated materials and providing new data.

0.2.3 The ECS SRIA and its position in the technology landscape

Electronics components and systems are key digital technologies enabling the development of numerous applications. As such, the ECS research and innovation priorities are significantly driven by application roadmaps and needs. To that effect, the **Key Application Areas** part of the ECS SRIA translates application roadmaps into requirements for ECS. Conversely, the **Foundational Technology Layers** part maps out future advances and potential new breakthroughs in applications. The ECS SRIA therefore promotes synergies with many neighbouring application-oriented communities. For example, the **Mobility** Chapter (3.1) has strong links with ERTRAC; the **Digital Industry** Chapter (3.3) with EFFRA; and the **Agrifood and Natural Resources** Chapter (3.5) with the working group of the Alliance for the Internet of Things Innovation (AIOTI) in Smart Farming and Food Security, and with Water Europe[14]. In each case, experts participated in the work of both groups. There are also close interactions and alignments with European PPP initiatives, such as 2Zero and CCAM, IHI, etc.

The **Cross-Sectional Technologies** part also leverages the links of the ECS community with other technology-oriented domains, such as the European Technology Platform for High Performance Computing (ETP4HPC), EuroHPC, the European Working Group on High-Performance RISC-V based reference processing architectures, and Big Data Value Association (BDVA), with strong relations with the **Edge Computing and Embedded Artificial Intelligence** Chapter (2.1). Likewise, the **Connectivity** Chapter (2.2) benefited from fruitful exchanges with the 5G Infrastructure Association, SNS and inputs from the European Cyber Security Organisation (ECSO), as reflected in Chapter 2.4.

Several contributors of the **Technology** parts are also actively involved in the elaboration of international roadmaps (e.g. the Heterogeneous Integration Roadmap (HIR)[15] in electronic packaging and integration, the IEEE International Roadmap for Devices and Systems (IRDS)[16] for the semiconductor industry), and the RISC-V Roadmap of RISC-V International[17] and the European Working Group[18].

To summarise, this ECS SRIA combines application-pull and technology-push with the objective of enhancing the fertile dialogue between technologists and technology users, and strives to include discussions of upcoming strategic value chains.

0.2.4 New and growing trends

Content wise, this revision was conducted with the goal to reflect the most recent technological and strategic trends of our industry. While already present in earlier editions, quantum technologies are getting additional attention in this new edition. The same holds for the irruption and disruption of AI tools in the practices of the ECS community, and the digital society at large. The importance of the RISC-V and Open-Source Hardware trend is further reiterated. Finally, the emergence of chiplets as a potential game-changer of our industry is discussed in this introduction, with several chapters going into the related technology requirements. The next pages provide some more details on the main changes of this edition.

[14] https://watereurope.eu/wp-content/ uploads/2019/07/Water-Europe-SIRA.pdf
[15] https://eps.ieee.org/technology/ heterogeneous-integration-roadmap/2019- edition.html
[16] https://irds.ieee.org/
[17] RISC-V-Introduction-_-Aug-2021.pptx (live.com)
[18] https://digital-strategy.ec.europa.eu/en/library/recommendations-and-roadmap-european-sovereignty-open-source-hardware-software-and-risc-v

A Chiplets

A chiplet is a tiny integrated circuit (IC) that contains a well-defined subset of functionality. It is designed to be combined with other chiplets in a "Lego-like" assembly[19].

In a sense, chiplets can be seen as a continuation of the trend already initiated with System-in-Package (SiP). With respect to System-on-Chip (SoC), the SiP approach allows to reduce the cost per square millimetre of silicon, since it is not required to fabricate all the circuit blocks using a process which could meet the performance criteria needed by all functions.

However, with respect to the monolithic SoC approach, SiP is impaired by limitations in bandwidth and latency for data transfer between its building blocks. Recent advances in heterogeneous integration technologies now allow to envision to have the “best of both worlds”: keep the advantages of SiP with circuit blocks built using, for each of them, the most adequate process technology in term of cost, power and performance, optimized for the chiplet particular function, while ensuring data transfer capabilities between those blocks close to what is currently achieved with a SoC. Other advantages of the chiplet approach (already present with SiP) are the possibility to reuse intellectual property and to test individual chips before assembly, improving the yield of the final device.

Chiplet technology has therefore the potential to lead to versatile and customizable modular chips, enabling reduced development timelines and costs. In addition, IP reuse improves design flexibility and efficiency. With all those promises, chiplets are sometimes touted as a revolutionary advancement in semiconductor technology[20]. It is especially relevant for Europe, as a path to deliver innovative solutions while keeping the dependence relative to the dominant, non-EU advanced CMOS manufacturers to a minimum.

While several vendors are now selling systems using chiplet technology, there is still many advances to be done before realising the vision of an open chiplet marketplace, including standardized interfaces, that would allow designers to select and assemble chiplets best suited for the system they want to design. In this SRIA, research focus areas relevant for the progress and enablement of the chiplet approach, as well as consequences of this technology trend, are mainly dealt with in chapter 1.2, but are mentioned in many other chapters, making it a cross-cutting theme of this document.

B Open-Source Hardware

Open-source hardware is hardware whose design is made publicly available so that anyone can study, modify, and further distribute the design, and/or freely make and sell hardware based on that design. The use of open-source hardware drastically lowers the barrier to design innovative SoCs. Indeed, this allows research centers and companies to focus their R&D effort on innovation, leveraging an ecosystem of pre-validated IP that can be freely

---

[19] Don Scansen, EE Times "Chiplets: A Short History », https://www.eetimes.com/chiplets-a-short-history/, consulted 25 July 2024

[20] Cadence PCB solutions, https://resources.pcb.cadence.com/blog/2023-all-about-chiplet-technology, consulted 25 July 2024

assembled, modified, and customized for specific applications, whereas currently the cost of design of a set of IPs in-house is available to only few companies. And the use of alternative 3rd party IPs licensed by companies imposes constraints on innovation due to architecture. Open-source is also a sovereignty tool and avoids licensing IPs from foreign third parties, in a geopolitical context characterised by trade wars and export control restrictions.

While open-source SW approaches are now in wide use, open source HW has really drawn attention beyond the academic circles only in recent years. In China, USA and India, large governmental funded programmes have been created in order to stimulate the use of RISC-V and Open-Source in industry and products. The widespread adoption of this approach in Europe would have many benefits.

Creating an ecosystem to support open-source hardware is essential as it cultivates innovation and aids dissemination. This can be first based on the implementation of a strategic "Governance Initiative'' to coordinate activities, maintain and promote the repository. Open-source projects only work well when they are attractive and provide innovation, but also when they are useful and trusted, so that at this point they can be self-supported by the community. One needs to make sure that knowledge is shared, accessible, maintained and supported on a long-term basis, as the success of open-source depends not only on the IP blocks but also on the documentation, support and maintenance provided. In particular, a one-stop-shop model with long-term activities and overall support (e.g., advice for licensing, productisation, etc.) can be promoted to help SMEs and start-ups.

The European strategy on open-source hardware should focus on application domains where there is a stronger impact, i.e. automotive, industrial automation, communications, health and aeronautics/defense. As these fields convey specific requirements (safety, security, reliability, power and communication efficiency…), the key differentiator between European core design efforts and existing players will be the attention given to the aforementioned requirements and also the extensibility with custom operations. Above all, a special focus should be placed on safety and security solutions, dealing with key aspects of collaboration, documentation, verification and certification in open-source communities.

A working group gathering European stakeholders was established in 2022. It drafted a technological roadmap, attached in Appendix A of this SRIA edition, starting with processors (RISC-V, beyond RISC-V, ultra-low power and high-end). The chiplet-based approach mentioned above is identified by the roadmap as a unique opportunity to leverage European technologies and foundries to create European HW accelerators interposers that could foster European More-than-More technology developments. Die-to-Die communication remains a missing link to leverage that approach, and its development in open-source is advocated by the working group.

In conclusion, the emergence of a European open-source hardware community should be encouraged and its initiatives be guided towards application domains where there is a strong impact. A technological roadmap has been drafted to build on the strengths of European industry and fill its gaps, while building blocks being developed will also need to be supported by the required tools, software supports, tests and documentation.

C Quantum Technologies

Disruptive technologies, where market positions are not established yet, are level-playing fields where Europe has all its chances to gain leadership. Quantum technologies fall in that category and therefore their evolution and maturity are closely monitored by the ECS SRIA expert community.

Quantum technologies are very diverse in nature, encompassing quantum computing and simulation, quantum communication, and quantum sensing and metrology[21]. As such, they go much beyond the scope of the ECS SRIA. Nevertheless, as identified in the European Quantum Flagship SRIA, the two domains overlap along two tracks:

- Technologies for dedicated quantum chips, required to fulfil the quantum challenges.
- Classical chips technologies for quantum (enabling technologies) required to support the industrialization and scaling of quantum technology.

In particular, the emerging field of quantum computing poses its own challenges for the ECS community in process technology, equipment, and materials:

- As there are still several candidates for becoming the standard quantum computing technology, a wide range of materials is relevant, together with innovations in process technology.
- New metrology capabilities are required, especially the measurement of electrical properties, such as local carrier mobility.
- To achieve practical applications, reliable fabrication, connection, and read-out of qubits need to be developed. The low temperatures at which most quantum systems are operated requires the development of cryogenic devices, to interface conventional electronics.

Beyond quantum computing, all quantum technologies related to sensing, communications and computing, including software, present significant challenges today relevant for the ECS SRIA. Given the level of maturity of those technologies, especially when it comes to quantum computing, those challenges are mainly discussed in the Long Term Vision chapter, but research focus areas derived from the needs of quantum technologies can also be found in other chapters of this document. This reflects that quantum technologies maturity is expected to evolve rapidly in the years to come, and that Europe should leverage its assets, with excellent research centres and a dynamic ecosystem of start-ups, to gain leadership in that promising field.

D Silicon Photonics

With the development of the global internet traffic, AI/ML and IoT, the demand for data centers and high-performance computing (HPC) is increasing. An extremely large link capacity for high-speed datacom interconnects between multi-cores or local/distant caches becomes compulsory. However, it is becoming increasingly difficult for the conventional electrical Cu

[21] Strategy Research and Industry Agenda, Published in February 2024 by the European Quantum Flagship

interconnects to meet such an ever-growing capacity requirement since they severely suffer from limited bandwidth and significant power consumption.

Silicon photonic integrated circuit (PIC), referred to as Si-photonics, is the technology that realizes photonic components and circuits on Si substrates using materials and fabrication process flows in the well-established microelectronic industry. Replacing electrons with photons to transmit data brings advantages including high-speed operation, low power consumption and high-capacity transmission. Building photonic integrated circuits (PICs) on a industry-standard Si platform promises additional merits, namely ultra-low cost, high-volume manufacturing, large integration density, advanced functionality and high scalability.

Furthermore, Si-photonics, with their CMOS compatibility and small size, weight, area, and power consumption, have the capability of accommodating growing volumes of computer data arriving in real time and at very high rates in large-scale distributed computing systems. The wavelength and spatial multiplexing properties of optics are beneficially translatable to executing critical bandwidth and communications-intensive connections between increasingly parallel computational resources.

Benefiting from an ideal Si/SiO_2 interface and a large refractive index contrast, Si is an ideal platform for implementing passive optical devices such as low-loss waveguides, mode converters, and multiplexers/demultiplexers. Although Si is a centrosymmetric crystal and, thus, not an ideal material for optical modulators, high-performance modulators have been demonstrated. Efficient Ge- and SiGe-based photodetectors operating in the telecom bands are also selectively grown on Si-photonics wafers.

However, if Si can efficiently transmit, modulate, and detect light, its indirect band structure precludes efficient light emission. The achievement of on-chip light sources, using Ge alloys for instance, represents a significant milestone in advancing the complete integration of silicon-based PICs. Therefore, III-V material (InP, GaAs) on-chip lasers are currently embedded in PICs using 2.5 and 3D heterogeneous integration. Monolithic integration of III-V materials is an attractive alternative due to its low cost and high integration density. However, it introduces a high density of crystalline defects, significantly degrading the laser's performance. Exploring methods of reducing these defects is paramount to advance monolithic integration toward high-density, large-scale silicon-photonic integration.

Similarly, for applications that require high-speed modulation, the properties of Si-based modulators are not sufficient. Hence, ferroelectric electro-optic materials such as $LiNbO_3$ and $BaTiO_3$ should be embedded in the PIC using heterogeneous integration (die-to-die, die-to-wafer, or wafer-to-wafer bonding).

The most established form of the technology is optimized for operation in the telecommunication wavelength bands near 1310 and 1550nm and has waveguides defined in the top Si device layer of a silicon-on-insulator (SOI) wafer. It is enabling power-efficient transceivers and large-scale PICs that address the demands of data communications, three-dimensional (3D) sensing, and computation accelerators. In addition to these established applications, there is a growing demand to develop silicon-photonic for new product applications that include chip-to-chip electrical/optical interconnect, automotive LiDAR, quantum computing, infrared imaging and biomedical and environmental sensing. The later applications require the development of PICs in the mid-IR region wavelength range (2–20µm) as it contains strong absorption signatures of many molecules.

E Artificial intelligence topics in the ECS SRIA

While not a totally new topic (it is used for example in several embedded systems for vision, enabling autonomous vehicles), artificial intelligence (AI) really came to the forefront for the public at large in 2023, with tools such as Chat-GPT (released for public use only on November 30th, 2022) and other generative AI (Dall-E, Stable Diffusion) that grabbed the headlines in mainstream media. The ECS Research and Innovation roadmap interacts with the development of AI tools and algorithms in two symmetrical ways: ECS are an enabler for new AI developments, and AI can in turn be a significant enabler for new ECS advances. Those two aspects are extensively discussed in this current edition of the SRIA.

- ECS as enabler of AI: Chapters 1.1 and 2.1, as well as the Long-Term Vision Chapter, stress the importance of moving the processing of AI algorithms locally on a hardware device (in deep-edge devices, edge devices or on-premise computing resources, depending on the application) close to where the data is generated (e.g. by a sensor), a trend known as “embedded AI” or “edge AI”. Benefits include reducing the energy consumption of the data infrastructure by transmitting only relevant data or pre-treated information, improving data protection, increasing security and resilience due to a reduced reliance on telecommunication links, reducing latency, and decreasing memory footprint. Many of the focus research areas identified in Major Challenge 1 (Advanced computing, in-memory, neuromorphic, photonic and quantum computing concepts) of Chapter 1.1 support this move towards AI at the edge. Likewise, the four Major Challenges (increasing energy efficiency, managing system complexity, increasing device lifespan and ensuring European sustainability of embedded intelligence), identified by Chapter 2.1, all cover research topics which will enable the development of a strong European embedded AI ecosystem. Energy consumption of AI solutions is an overarching issue due to the expected widespread increase of their usage, and it is essential to explore new concepts and architectures (bio-inspired and other ones) to respond to that concern, which could turn out to be a major roadblock if not addressed properly.
- AI as enabler of ECS: this aspect is discussed in Chapter 2.3, where several examples of AI use are listed to tackle Major Challenge 3 (managing complexity), such as AI-based methods for the architecture exploration and optimisation, to achieve a global optimum, AI-based guidance in the V&V process, and automatic generation of test cases with AI support. Moreover, in its entirely new section devoted to Machine Learning and Artificial Intelligence, the Long-Term Vision Chapter states that AI/ML methods are increasingly adopted in design space exploration at several stages of circuit design, as well as in design testing and verification. The potential of these AI-based methods is not limited to circuit design and extend to large-scale ECS products encompassing HW and SW, as well as multi-physical, distributed systems. They are also expected to provide guidance to engineers in multi-risk (safety, security, privacy,

and other trustworthiness risks) optimisation problems. Gathering / generating the appropriate data to train the models is identified as one of the most difficult issues to be addressed.

A third aspect is that AI adds complexity to the systems it empowers. This in turn raises additional issues at the design phase, as addressed by Chapter 2.3. There, Major Challenge 1 (Enabling cost- and effort-efficient Design and Validation Frameworks for High Quality ECS) stresses the need to develop new design and verification and validation methods for AI-based ECS, including ECS evolving during lifetime.

Finally, a purely technical approach to develop trustworthy and explainable AI, or having AI in the loop with verification of its results might not be enough to address the growing public concerns about the potential dangers of AI (pushing many governments across the world to sign the Bletchley Declaration on AI safety). In the years to come, cooperation between multidisciplinary partners with backgrounds in AI, ECS and social sciences, as well as access to "foundation models" by European academia and industry will be required to achieve and ensure the acceptance of efficient and responsible AI-based ECS, reflecting European values.

F Ensure engineering support across the entire lifecycle of complex ECS-based systems

Modern ECS-based systems are complex, and they cannot be properly elaborated and used without the appropriate engineering support across the entire lifecycle, from requirements analysis to design, development, deployment/commissioning, operation/management, remote-maintenance repair and overhaul, retirement/recycling and evolution.
Engineering support represents a key factor for achieving the four Main Objectives as it:

- impacts industrial competitiveness by simplifying life cycle management, and improves the quality of the engineering process, making it more cost-effective and agile.
- simplifies and improves the development of trustworthy ECS technologies, products and applications.
- supports sustainability and resilience that reduce lifecycle management costs, as well as ensuring the automation and continuity of operations.
- is fundamental to unleashing the full potential of intelligent and autonomous ECS, which requires completely new approaches to engineering, design and development methodologies, as well as toolchains and tools.
- improves professional training and education by strengthening and developing new and specific skills.

ECS engineering plays a key role in Pillar 1 of the Chips Act, which includes specific instruments to support the engineering process, specifically in the critical area of new chips design, to increase design activities and create a strong design ecosystem in Europe. A Virtual Design Platform (VDP), providing design tools, an IP library for various technologies, and PDKs to access the fabs, will be established to simplify and accelerate the process of going "from the lab to the fab", and make that process more accessible for EU companies, start-ups and SMEs, especially fabless. Specific Design Enablement Teams (DET) will support this process providing targeted assistance on designing chips and more complex systems, with the presence of experts in the design flow, foundries, and pilot lines, from virtual prototyping via tape-out and engineering samples to high volume production. To this regard, Competence Centres will be established to address the skills challenges that Europe is experiencing and provide education and professional training solutions for the next generation of chip designers: DET and a Competence Centre will provide advice, training, and skills-building solutions, supported by the industry-grade training framework offered by the Virtual Design Platform. Chapter 2.3 and other chapters address many of the RD&I challenges that the VDP and Competence Centres will have to face in the next decade.

0.2.5 Further changes vs. previous edition

On top of these general trends, individual chapters also reflected the latest evolutions: The respective scopes of Chapter 1.1, Process Technology, Equipment, Materials & Manufacturing, and 1.2, Components, Modules & Systems Integration, were refined to have clear boundaries: The former chapter covers all process technologies, equipment and materials' research and innovation to enable CMOS compatible semiconductor chip and packaged chip manufacturing inside a cleanroom environment, as long as the wafer is not diced, while the latter one covers packaging, integration on board and module level. Text and Major Challenges in these two chapters have been revisited accordingly, including the addition in chapter 1.1 of a Major Challenge on Advanced packaging, assembly and test equipment solutions. That chapter also added a MC on "Sustainable semiconductor manufacturing". Chapter 1.3, Embedded Software and Beyond, addresses needs deriving from the advent of software defined vehicles, requiring collaborative embedded SW development processes and toolchains. New concepts for programming languages, such as Rust, have been introduced, and software engineering practices in the AI-assisted engineering era are discussed. Updates in lifecycle management and hardware virtualization for efficient SW engineering were also introduced. In Chapter 1.4, System of Systems, the two Major Challenges "Control in SoS composed of embedded and cyber-physical systems" and "SoS monitoring and management" have been merged.

Regarding Cross Sectional chapters, Chapter 2.1, Edge Computing & Embedded AI, has been heavily restructured, to regroup the two formerly separate sections on Edge Computing and Embedded AI, reflecting the convergence of the two domains. Major trends such as complexity management utilising AI and the decomposition of complex SoCs into chiplets and interposers are discussed in details. Chapter 2.2, Connectivity, was edited to reflect the decisions taken during ITU World Radiocommunication Conference 2023 (*WRC-23*)[22], as well as the latest thoughts of the industry regarding the use of the various frequency bands when deploying 6G. In Chapter 2.3, Architecture and Design: Methods and Tools, the two former Major Challenges "Extending Development Processes and Frameworks" and "Managing new functionality in safe, secure and trustworthy systems" were merged into one, "Enabling cost- and effort-efficient Design and Validation of High Quality ECS", while a new MC was introduced on "Enabling Sustainable Design for Sustainability", with associated new research focus areas.

[22] This conference took place in Dubai, United Arab Emirates, from 20 November to 15 December 2023

Among the Application chapters, in Chapter 3.1, Mobility, the challenges regarding the software-defined vehicle have been expanded into two Major Challenges, "SDV hardware platforms: modular, scalable, flexible, safe & secure" and "SW platforms for SDV of the future: modular, scalable, re-usable, flexible, safe & secure, supporting edge2cloud applications", while a new Major Challenge has been introduced on "Edge2cloud mobility applications: added end-user value in mobility". The multimodal mobility topic has been moved to Chapter 3.6, Digital Society. Chapter 3.2, Energy, insists on the challenge of meeting significant energy demand by new technologies and their applications, used by an interconnected society and based on the availability of an infinite number of accessible data – AI, generative AI, large language models, crypto currency and the exponential use of them in our connected world. This challenge comes in addition to the targets set by the European Union for a renewable energy share of 32 percent and a Greenhouse gas emission reduction of 55 percent by 2030. Chapter 3.4, Health and Wellbeing, is putting more focus on demographic change as one of the major drivers (and listing of possible technical solutions to cope with it). The update of Chapter 3.5, Agrifood and Natural Resources, develops the themes of climate-smart agriculture, wearable plant sensors, carbon sequestration and agriculture passport. In Chapter 3.6, Digital Society, more attention is paid to the use of AI-based tools (such as ChatGPT and Claude). Measures against fake video and audio are included, as well as the increasing importance of cybersecurity and trustworthiness. Furthermore, diagrammes have been updated with more recent versions and post-Covid aspects have been made less prevailing.

0.2.6 The ECS global timeline for Europe

The 2025 ECS SRIA lists a number of milestones to be reached in the short term (2025–2029), medium term (2030–2034) and long term (2035 and beyond) via collaborative research projects across Europe, reflecting the ambition of the ECS industry towards the achievement of the four Main Objectives identified above.
The following figures summarise the most salient milestones to be reached in the various domains covered by the ECS SRIA over the three time periods:

- **Short term (2025–2029)**
  The industry has a precise idea of what will be achieved during this short-term timeframe.
- **Medium term (2030–2034)**
  There is already reasonably good knowledge of what can possibly be achieved.
- **Long term (2035 and beyond)**
  Expected achievements are more of a prospective nature.

Including a milestone in each of these time periods means that the described features are expected to be available at TRL levels 8–9 (prototype or early commercialisation) within that timeframe. For example, the **Components, Modules and Systems Integration** Chapter expects that, within the next five years (short term), the materials that enable recycling and repair will be available. These materials will allow for the deployment of the monitoring of forests, fields and oceans, as envisioned by the **Agrifood and Natural Resources** Chapter over the same time horizon. In parallel, this monitoring will gain in efficiency due to the development of advanced AI edge solutions leveraging open source or alternative strategies, as forecast by the Chapter on **Edge Computing and Embedded Artificial Intelligence**.

The above example also clearly shows that progresses in the various domains covered by the ECS SRIA are deeply interconnected. Innovation in one area is building upon, or being driven by, innovation in other areas. Similar examples could, of course, be developed for the other time horizons, as represented in Figure 0.6, 0.7 and 0.8. More detailed diagrams, including additional milestones, are presented in the individual chapters.

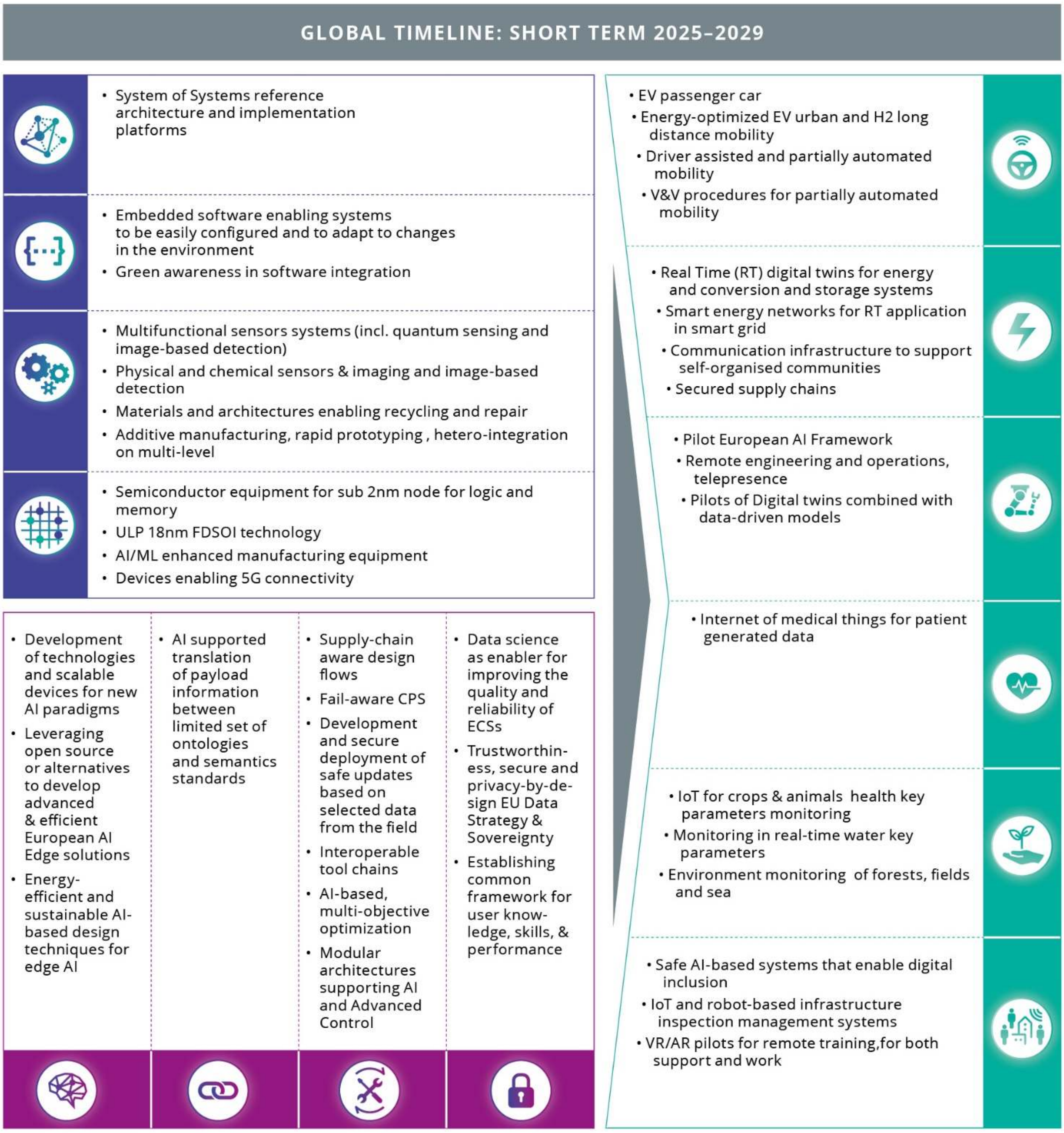


Figure 0.6 Global Timeline: Short term 2025–2029

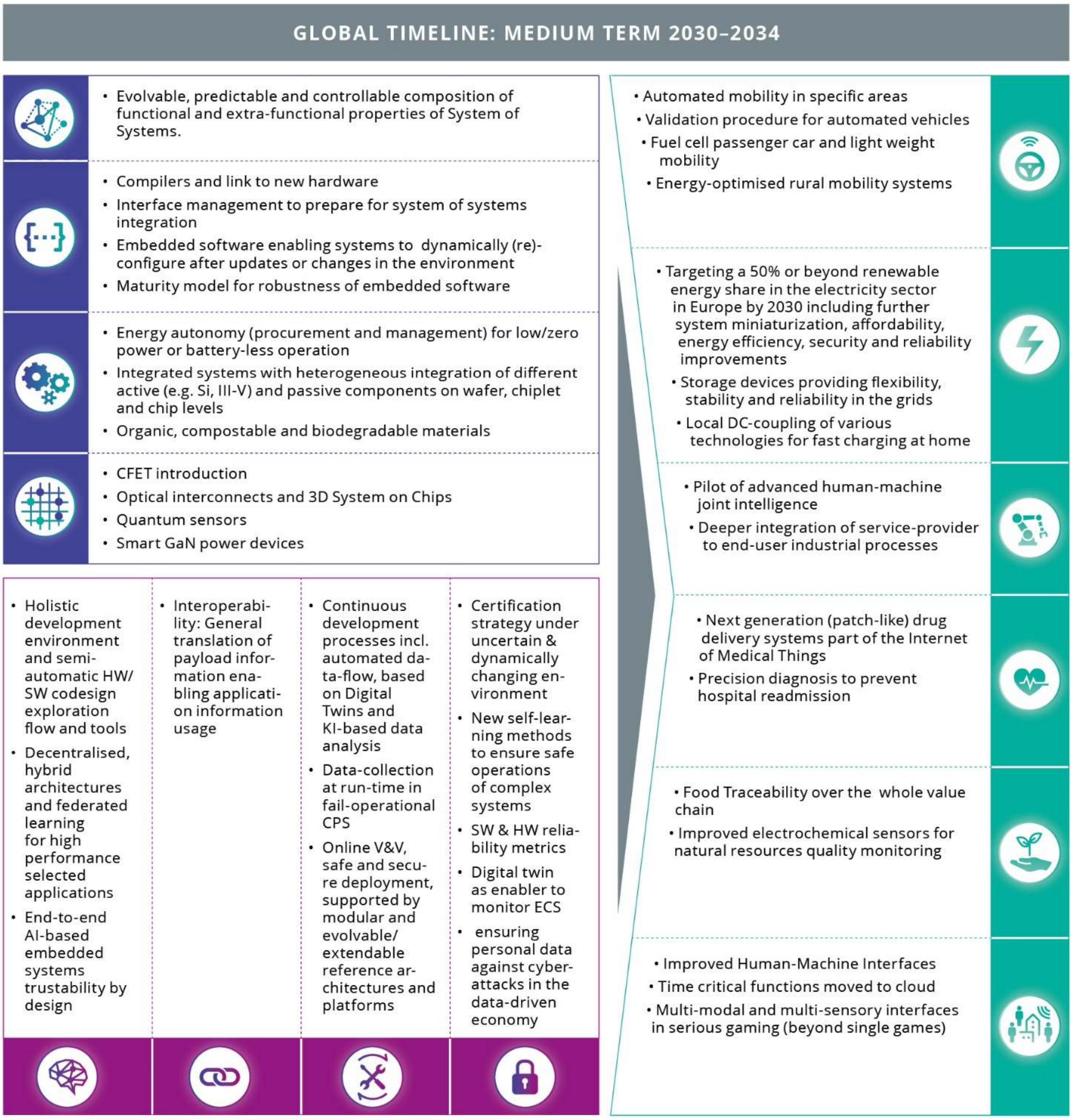


Figure 0.7 Global Timeline: Medium term 2030–2034

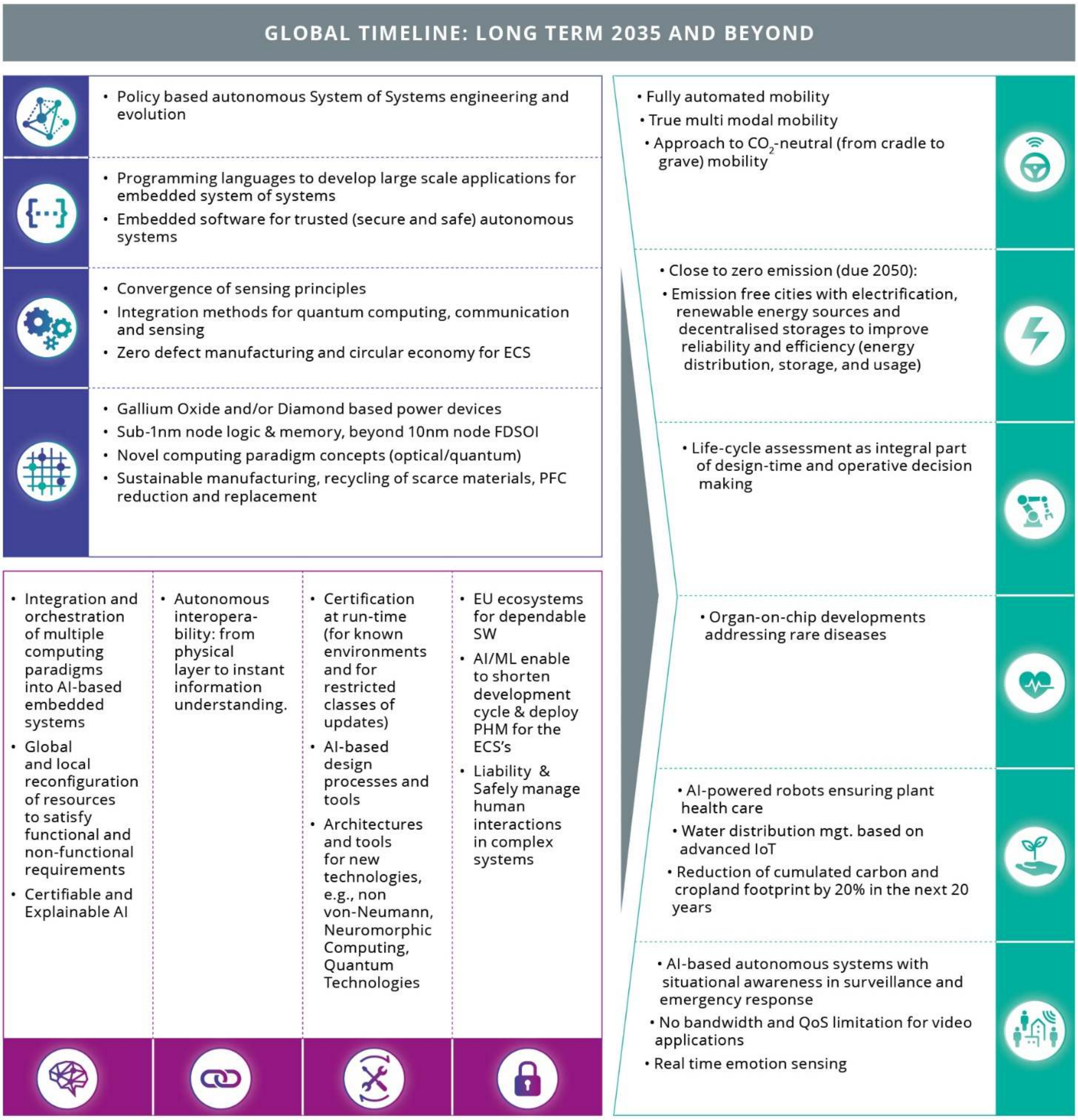


Figure 0.8 Global Timeline: Long term 2035 and beyond

0.2.7 The ECS SRIA outline

The following diagram provides an outline of the entire ECS SRIA to clarify the roles of the chapters, the technology domains they cover and the synergies between them, simplifying the comprehension of the ECS SRIA and its "navigation".

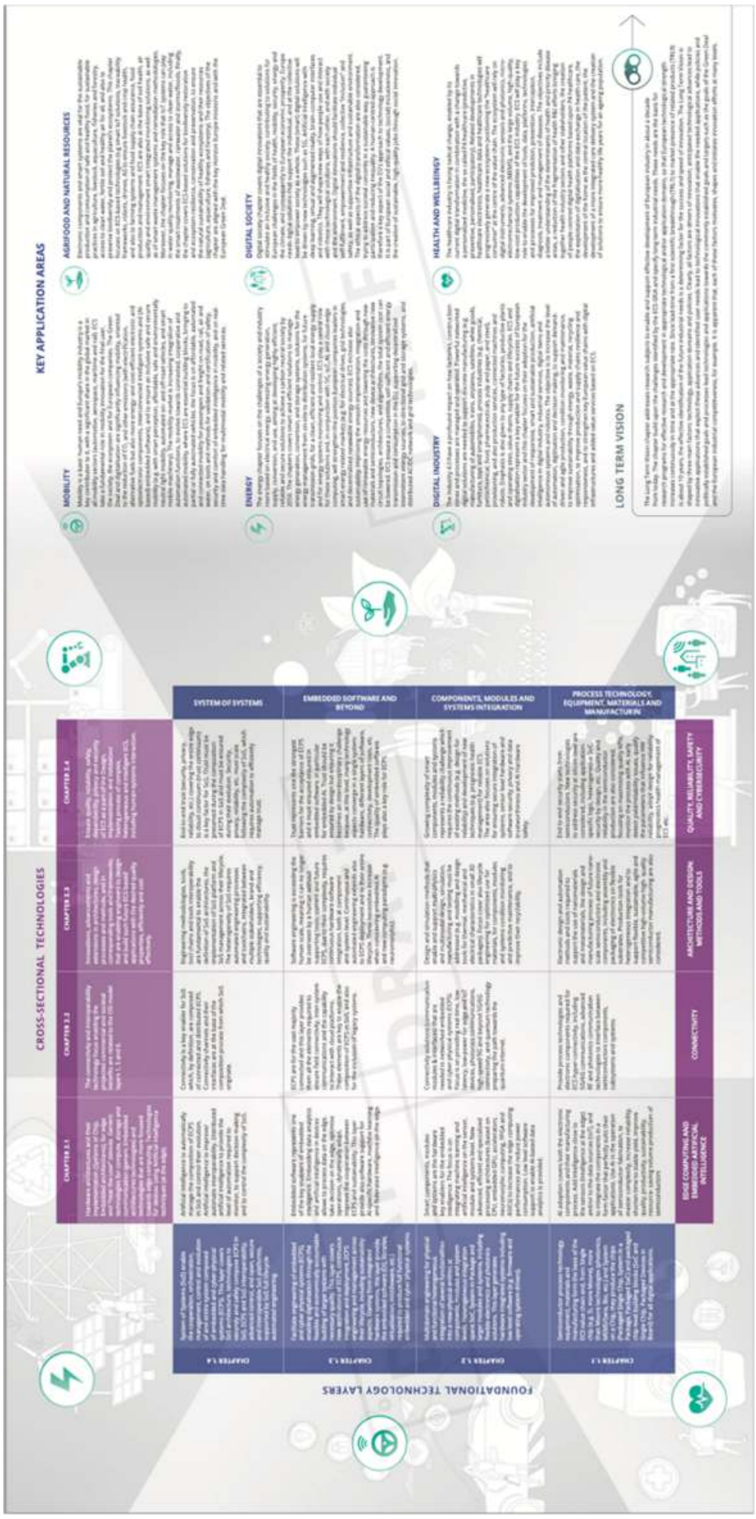


Figure 0.9 SRIA outline

0.3. PART 3 - HOW ?

0.3.1 Navigating the ECS SRIA

As mentioned, the ECS covered by this document is very wide-ranging, and involves many technical disciplines in materials, processes, hardware and software. Therefore, most readers will only want to read those chapters that cover the disciplines of their interest, without being willing to go through all the technical details presented in the ECS SRIA.

The structure of all the **Foundational Technology** and **Cross-Sectional** chapters is identical. In the **Key Application Areas** chapters, the authors explore each domain with the application demands as the main focus, not the technical challenges.

A **Glossary** describing the terms used in the document, as well as a List of **Acronyms** used in the document, can be found in the **Appendix**. The Analytical Index available in earlier releases of the ECS SRIA is now substituted by an advanced Search functionality. At the end, the reader can also find a **List of Contributors** who collectively wrote this ECS SRIA.

Finally, to highlight the synergies/links between the chapters and provide hints to the reader, cross-references have been introduced alongside the text.

Cross-references consist of the Chapter icon and appear alongside the text. When hovered on, they indicate the relevant chapter number and the topic or concept described in the text is also highlighted. When clicked on, they will navigate you to the referenced Chapter.

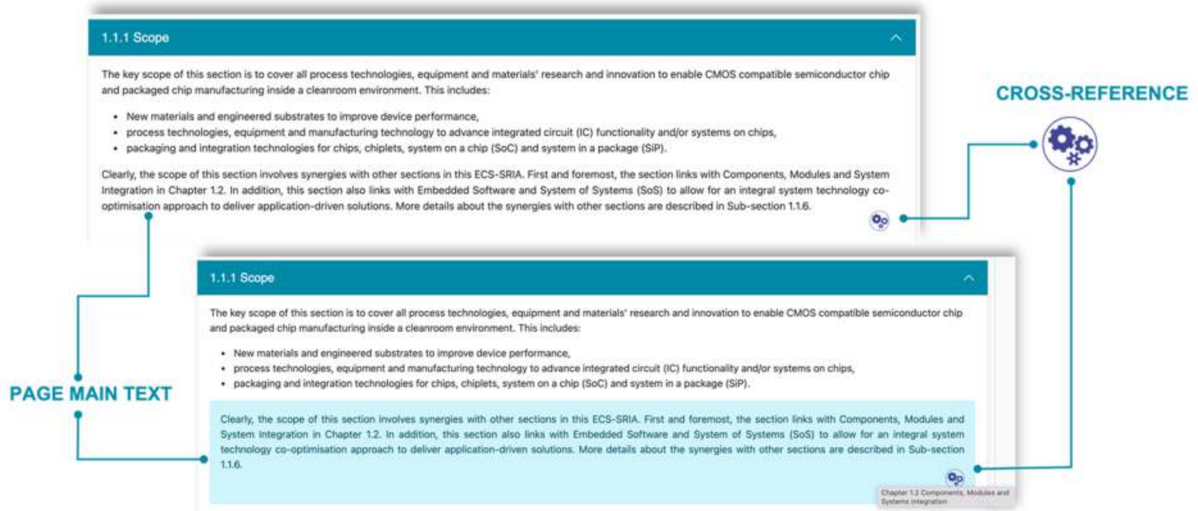


Figure 0.10 Cross Reference

References within a text are indicated by small numerical markers embedded in the content, e.g. [0]. Upon clicking one of these markers, readers are instantly directed to the corresponding entry within the References section. This feature allows readers to seamlessly access the sources from which the information has been derived.

Furthermore, by clicking on the same footnote number located in the **References** section, readers can effortlessly return to the exact point within the main text where the reference was initially cited. This interactive function not only enhances the reader's comprehension but also streamlines the process of verifying and exploring the sources behind the presented information.

0.3.2 Make it happen

Whatever their scientific and technological excellence, research projects can achieve significant societal and economic impact for the EU only if several “innovation accelerators” are in place, which will bring the research results to market. Standardisation and regulation, education and training, international cooperation, and research infrastructures are a few of these accelerators.

Discussing those topics is beyond the scope of the ECS SRIA, which is the expression by the industry of its Research and Innovation vision. In essence, the SRIA is designed to be funding instrument agnostic, so that it can constitute the basis of calls for various programmes, such as Horizon Europe, Eureka, or national initiatives. That being said, some of the accelerators listed above are being set up in the frame of the Chips Act. In particular, the Design Platform and the Pilot Lines included in the Chips for Europe Initiative have the objective of supporting and accelerating technological capacity building and innovation in the Union by bridging the gap between the Union’s advanced research and innovation capabilities and their industrial exploitation.

While it is outside of its scope to indicate how these research infrastructures and services should be run, the ECS SRIA covers the technological challenges to be addressed, from the Design Platform, via the different Pilot Lines towards accelerator projects to speed up the transfer from the lab to the fab. It is therefore important to ensure that their implementation most effectively supports the research needs exposed in the SRIA. To that effect, the following table identifies, for the Design Platform and for the Pilot Lines identified as of the time of writing of this SRIA edition, research topics which are expected to be supported by those instruments.

	Design platform	Advanced 2nm and beyond	FD-SOI
1.1 Process Technology, Equipment, Materials and Manufacturing		Launching ground for new processes, equipment technologies and materials	Low-power consumption, radiation hardness, More than Moore app.
1.2 Components, Modules and Systems Integration	Improve design capabilites to become a closed loop (i.e., to include feedbacks from the production process and from the field use, respectively) as well as define the new sets of interfaces for the complex integration solutions at die / module / system levels as needed for implementing heterogeneous and chiplet approaches - in particular for ECS applications that will be exposed to demanding and harsh environments (as these ECS are essential for our European backbone industry -automotive, energy, industry, health, ...- and not sufficiently and securely addressed by the worldwide leading players).	Impact of advanced node inflections like backside power distribution networks, forksheet, CFET and 2D material channels 3D heterogeneous integration in chiplet implementation	Provision of advanced logic building blocks with advantageous trade-offs in terms of speed, power consumption, robustness with respect technological simplicity and cost
1.3 Embedded software and Beyond		Design Technology Co-Optimisation	Design Technology Co-Optimisation
1.4 System of Systems		System Technology Co-Optimisation	System Technology Co-Optimisation
2.1 Edge Computing and Embedded Artificial Intelligence	Provide non-differentiating IPs (I/Os, memory interfaces, etc) Support the Open Source Hardware community in Europe Provide tools for embedded AI, such as tools allowing to quantize and decrease the size of Neural Networks for embedded accelerators	Drive PPACE (Power - Performance - Area - Cost - Efficiency) improvements for advanced nodes to increase the energy efficiency of computing systems Research PDK's to support managing the increasing complexity of systems	Energy efficiency, embedded non-volatile memories
2.2 Connectivity			* Enable the development of power efficient connectivity solution leveraging European-based semiconductor technology * Enable the development of innovative connectivity solution at mmW and THz frequencies
2.3 Architecture and Design : Methods and Tools	Support EDA research and innovation, in particular: - Exponentially increasing design complexity (MC 3) and Increasing Diversity (MC 4) - Sustainability (MC1) - Emerging technologies (MC1) - Increased automation and operability (MC1, MC2, MC3) - Multidisciplinary design (MC4)		Automated application of back-biasing for adjusting power and performance
2.4 Quality, Reliability, Safety and Cybersecurity	Widen the implementation of Design for X (DfX) besides the design for functionality and performance - in particular for the heterogeneous and chiplet implementations to be used in demanding and harsh environments. Implement 'closed loop design' approaches (= with feedbacks from manufacturing, testing, ... field use) X ... manufacturing, testing, reliabiltiy, safety, security, repairability, disassembly+reuse, sustainability, ... Chapter 2.4 highlights the importance of secure design environments to protect against cyber threats and secure and reliable environment both in terms of robust cybersecurity measures and safety properties that are needed due to the complexity and interdependencies of modern ECS. Chips JU Design platform aims to provide secure, cloud-based design facilities, integrating tools and support services to enhance cybersecurity. New cybersecurity measures and technologies implemented in a controlled environment before bringing them in real-world systems. This could include testing new encryption algorithms, security protocols, or methods for protecting privacy, enhancing resilience of electronic systems against cyber threats and fostering a safer digital ecosystem across the EU.	Design and integration, is crucial for developing advanced semiconductors. The pilot line can benefit from insights on how to optimize system architecture, integrate components, and ensure efficient communication between different parts of the chip. Security and Privacy Considerations: The pilot line should incorporate robust security measures to protect chips from cyber threats. In addition, it should establish rigorous testing protocols to ensure the functionality, reliability, safety and security of the developed chips including vulnerability assessments and penetration testing.	

	Design platform	Advanced 2nm and beyond	FD-SOI
3.1 Mobility	High Performance automotive SOC for Software Defined Vehicles based on RISC-V (including necessary accelerators and support for automotive bus systems as well as automotive trust-ability concepts)		
3.2 Energy	Energy efficiency is one of the major factors to reduce energy consumption - design for energy efficient devices, process technologies for energy efficient operation and integration and packaging technologies for all the advanced technologies, either logic or power electronics. To manage the demand of the interlinked society for more and more communication and in parallel a sustainable way of operation including fulfilling the transition towards the Green Deal objectives plenty of control and forecasting systems will be required in addition to highly miniaturized, safe, connected and efficient power conversion and distribution systems.		
3.3 Digital Industry			
3.4 Health			
3.5 Agrifood and Natural Resources	New advanced circuits needed to develop innovative and cost-effective solutions related to the chapter challenges could be designed (through the Design platform programme) and fabricated (through the Pilot lines programme). Examples include sensors for water quality monitoring, GHG emission measurement, plant and soil health.		In agriculture, low cost, highly energy-efficient and self-contained components are essential. They include highly innovative types of sensors related to agriculture and intended to preserve natural resources, highly efficient processors including AI capabilities and long-range RF solutions.
3.6 Digital Society			
Long Term Vision		Foundational studies for new processes, equipment technologies and materials	

	Advanced Packaging and Heterogeneous Integration	Wide Band Gap	Integrated Photonics	Quantum technologies	Other pilot lines
1.1 Process Technology, Equipment, Materials and Manufacturing	Introduce materials and process innovations as well as advanced manufacturing, test and inspection equipment for future AP/HI systems.				
1.2 Components, Modules and Systems Integration	Enable enhanced and diversified functionalities (e.g. combined sensing, processing, communication, ...) in small form factor electronic components and systems. Chapter 1.2 is the natural playground for the reunion of the components provided by the other PLs by means of the advanced heterogeneous system integration enabled by this one.	Provision of electronic components and modules made around alternative semiconductors that allow improved systems for energy/power management or operation in harsh (thermal, electric, radiation…) environments	Provision of photonic components and modules that on their own or hybridized with electronics either enable physical-digital interaction of actionable systems by means of light or boost the performance of digital communications	Mature quantum devices to be integrated in high performance/high sensitivity/highresolution sensor systems, and breakthroughs in quantum comunications and quantum computing	Platforms leading to the scalability of elements enabling connection between the digital and physical worlds (e.g. MEMS, integrated photonics, power electronics, quantum approaches…) in silicon or silicon alternative technologies will provide, together with logical circuitry, additional essential building blocks to be integrated in full fledged electronic systems
1.3 Embedded software and Beyond					
1.4 System of Systems					
2.1 Edge Computing and Embedded Artificial Intelligence	Advanced packaging technologies allow the integration of diverse and specialized components such as processors, GPU/FPGA, memory, sensors and communication chips in a single package. This enables new, more powerful and energy-efficient edge computing, AI and communication solutions in small form factors.				
2.2 Connectivity	* Enable a European ecosystem that can support heterogeneous integration (multi-die system in a package, advanced assembly capability, advanced substrate manufacturing, etc.) to help European players capture higher value in the connectivity market. * Enable the development of innovative Antenna in package solution at mm-wave and THz frequencies * Enable a sovereign European packaging ecosystem to secure the supply chain of European semiconductor players (especially in key areas such as space were required manufacturing scale limits the possibility to have access to Asian OSAT)				
2.3 Architecture and Design : Methods and Tools	Proving ground for new design methods mastering the differences in materials, technologies, and processes that are heterointegrated into a single package.				

	Advanced Packaging and Heterogeneous Integration	Wide Band Gap	Integrated Photonics	Quantum technologies	Other pilot lines
2.4 Quality, Reliability, Safety and Cybersecurity	Develop new processes ensuring quality, reliability and safety of hetero-integrated chips and systems. Heterogeneous integration involves connecting diverse components (e.g., CPUs, memory, sensors) within a single package. interconnect challenges are identified in this chapter such as signal integrity and reliability. The pilot line should explore innovative interconnect solutions. The pilot line could address also robustness via testing methodologies, failure analysis, and reliability prediction defined in the chapter. This includes stress testing and fault tolerance.	Developing efficient power devices requires careful design and integration of device architecture to optimize performances as well as guidance on designing RF circuits, minimizing losses, and achieving high-speed performances.	Photonic Integrated Circuits (PIC) development involves novel system designs and explore integration techniques for miniaturize and consolidate optical elements, similarly to how electronic components are integrated on microchips in traditional integrated circuits. The pilot line should supports breakthrough technologies like quantum computing, AI, and neuromorphic systems as well as functionalities such as security and energy efficiency to be considered during development.	Quantum Computing computers leverages on quantum bits (qubits) to perform complex calculations exponentially faster than classical computers. this implies improving qubit stability, error correction, and scalability. Links with the chapter are identified on optimizing the architecture of quantum computers in terms of efficient integration of quantum components. Quantum Communication/criptography ensures secure transmission of information using quantum key distribution (QKD), exploring quantum encryption, and entanglement-based communication. Links with chapter are: communication protocols and encryption techniques Quantum Sensing measuring physical quantities such as time, acceleration, and magnetic fields. Applications impacted include navigation, imaging, and environmental monitoring. Link with chapter are sensors to ensure quality and reliability. Quantum Materials and Devices materials (e.g., superconductors, topological insulators) and develop quantum devices (e.g., single-photon detectors, quantum memories). Links with the chapter are material properties that could be relevant to quantum technologies Quantum Algorithms and Software algorithms and tools for solving real-world problems. links with chapter are in terms of algorithm approaches and software development approaches.	

	Advanced Packaging and Heterogeneous Integration	Wide Band Gap	Integrated Photonics	Quantum technologies	Other pilot lines
3.1 Mobility	New generation of environmental sensors (or combined sensors), which simplify and improve object and lane detection, work in difficult (severe) weather conditions and situations (as tunnel exit). Can be supported via Heterogenous integration pilot line. Chiplets will allow for higher functionality and future capabilities to combine processing, sensing, and memory out of different nodes		New generation of environmental sensors (or combined sensors), which simplify and improve object and lane detection, work in difficult (severe) weather conditions and situations (as tunnel exit). Can be supported via Heterogenous integration pilot line, but also via other PL on photonics, quantum sensors.		
3.2 Energy	Energy efficiency is one of the major factors to reduce energy consumption - design for energy efficient devices, process technologies for energy efficient operation and integration and packaging technologies for all the advanced technologies, either logic or power electronics. To manage the demand of the interlinked society for more and more communication and in parallel a sustainable way of operation including fulfilling the transition towards the Green Deal objectives plenty of control and forecasting systems will be required in addition to highly miniaturized, safe, connected and efficient power conversion and distribution systems.				
3.3 Digital Industry	Facilitate advanced, cost- and energy-efficient integrated systems for key application areas.				
3.4 Health					
3.5 Agrifood and Natural Resources	In agricultural devices with a minimal form factor, low power consumption and low cost are essential. The wide variety of devices, related to the multiple applications, requires modular and heterogeneous integration including multiple sensors, advanced processors and multiple RF protocols.				
3.6 Digital Society	Facilitate advanced, cost- and energy-efficient integrated systems for key application areas.				
Long Term Vision	Establish and nurture advanced packaging and heterointegration expertise in EU. Facilitate and promote AP/HI technology and infrastructure access to foster electronics device and systems innovations.				

The Design Platform, as proposed in the European Chips Act, should also support research in several cross-sectional technologies:

- On the one hand, the development of edge AI and embedded computing chips (Chapter 2.1) will be facilitated if the Design Platform covers the following aspects:
    - Providing as many non-differentiating IPs (for instance I/O's, memory and communication interfaces, etc.) as possible, allowing to have a one-shop entry for start-up/SMEs and academia to validate into silicon their new architecture ideas in the field of accelerator for IA (at the edge) and embedded systems.
    - Supporting the Open-Source Hardware community in Europe, where the Design Platform could be linked with Open-Source repositories and allowing access of the instances (and the tools to use them, such as compilers, OS and basic middleware).
    - One specific topic that the Design Platform should add is to give access to the tools specific for embedded AI, such as tools allowing to quantise and decrease the size of Neural Networks for embedded accelerators (and perhaps to learning databases) so that it will be a single entry for using all those tools in a coherent environment. That will imply certainly to bridge to other platforms either from European projects (NeuroKit2E for example), but perhaps also with non-European repositories such as HuggingFace.

- On the other hand, the Design Platform could be used to investigate many EDA challenges identified in Chapter 2.3, such as:
    - Exponentially increasing design complexity (Major Challenge 3) and diversity (Major Challenge 4), including multidisciplinary design.
    - Sustainability (Major Challenge 2)
    - Support for emerging technologies: quantum computing, neuromorphic, edge AI, …

- Increased automation and interoperability in the design flow, including for multi-vendor solutions

In a nutshell, the Design Platform and the Pilot Lines included in the Chips for Europe Initiative have the potential to offer a valuable research, development and testing ground, a learning environment to advance innovation, enhance manufacturing capabilities, and accelerate the development of cutting-edge electronic products. To make that promise come true, it is of utmost importance that their research roadmaps are established in strong synergy with the contents of the ECS SRIA and that the results are prepared for transfer into volume manufacturing.

The ECS community is positive toward those two new instruments under the pre-condition of appropriate involvement of industrial stakeholders in advisory bodies of these two instruments. For the Pilot lines the projects have to be designed to explore new technologies with clear outcome for the European industry at large, and with feedback mechanisms ensuring this is implemented and established over time. For the Design Platform the rules of access and use of the resulting designs (Silicon IP) or chips for companies or laboratories will have to be clearly and carefully crafted to avoid competition distortion.

1

Strategic Research and Innovation Agenda 2025

FOUNDATIONAL TECHNOLOGY LAYERS

1.1

Foundational Technology Layers

PROCESS TECHNOLOGIES, EQUIPMENT, MATERIALS AND MANUFACTURING

1.1 PROCESS TECHNOLOGY, EQUIPMENT, MATERIALS AND MANUFACTURING

Semiconductor process technology, equipment, materials and manufacturing form the foundation of the ECS value chain producing the chip and packaged chip-level building blocks for all digital applications.

Nano- and microelectronics are key to achieving digital sovereignty in Europe, and they offer a range of solutions for a green and sustainable society. If Europe wants to control the development of a digital future fitted to its citizens and their requirements, as well as its social, economic, industrial and environmental goals, it needs continuous innovation in the field of semiconductor technology.

1.1.1 Scope

The key scope of this section is to cover all process technologies, equipment and materials' research and innovation to enable semiconductor IC manufacturing inside a cleanroom environment. This includes:

- New materials and engineered substrates to improve IC performance,
- Process technologies, equipment and manufacturing technology to advance integrated circuit (IC) functionality and/or systems on chips,
- Wafer level integration addressing a technology of packaging a die while it is still on the undiced wafer or bonded/attached to a wafer (D2W). Protective layers and electrical connections are added to the substrate before dicing.

Clearly, the scope of this section as indicated in figure 1.1.1 below involves synergies with other sections in this ECS-SRIA. First and foremost, the section links with Components, Modules and System Integration in Chapter 1.2. In addition, this section also links with Embedded Software and System of Systems (SoS) to allow for an integral system technology co-optimisation approach to deliver application-driven solutions. More details about the synergies with other sections are described in Sub-section 1.1.6.

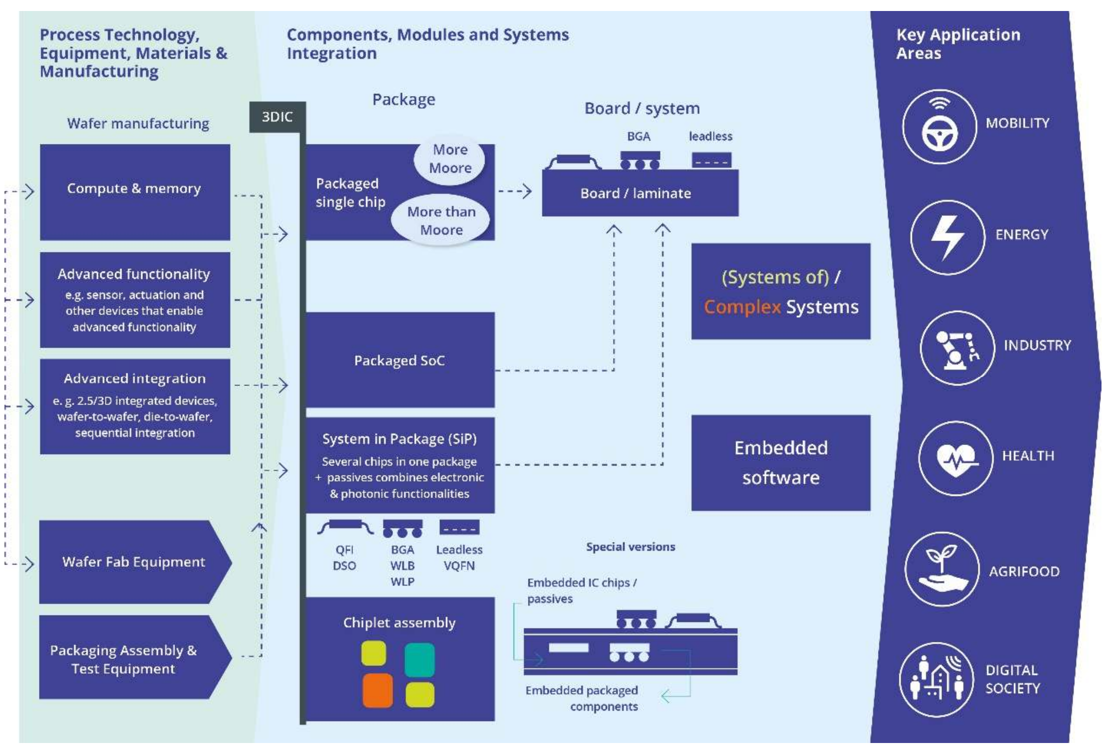


Figure 1.1.1 - The chip and packaged chip-level building blocks are the starting point for the other ECS-SRIA Chapters

1.1.2 Application breakthroughs

The main breakthrough enabled by the technological advances discussed in this section concerns the reduction of energy consumption in the various electronic components without any decrease in their performance.

In 2022, the globally consumed power of data centres alone was 240-340 TWh[1], which represents between 1-1.3% of global final electricity demand. This excludes energy used for cryptocurrency mining, which was estimated to be around 110 TWh in 2022, accounting for 0.4% of annual global electricity demand. Investing in more efficient IT hardware including microchips will provide the means to flatten this curve whilst data centre workloads are expected to dramatically increase[2].

Reducing the energy consumption of electronic components is essential for improving the autonomy of electric and hybrid vehicles, the lifetime of battery-powered sensors (for health monitoring, preserving natural resources such as water through more efficient irrigation, etc.), as well as for the development of autonomous sensors with energy harvesters and energy storage.

Since moving data from the logic cores to the adjacent memories is the main contributor to the energy consumption of logic devices (microprocessing units (MPUs), microcontroller units (MCUs), etc), their conventional von Neumann architecture must be drastically changed in close co-optimisation with other technology innovations. Near-memory or in-memory computing and neuromorphic computing are new architecture paradigms that strongly reduce the movement of data, and accordingly allow decreased overall energy consumption. Specific low- power transistors, memory and 3D-integration technologies need to be developed to ensure close coupling between computer and memory blocks.

The adoption of wide bandgap materials such as GaN and SiC is crucial for allowing higher operating temperatures and reducing the switching losses in power electronics for a broad range of power systems, such as smart phone/tablet chargers, industrial power supplies, power supplies for servers, etc., and - very important -, electric vehicles, as well as to increase their range. GaN/SiC is also important for increasing the power efficiency of 5G RF base stations. In addition, GaN/Si and GaN/SOI can induce the same effect in RF front-end modules when combined with high thermally conductive materials.

The exponential increase in internet traffic (with a CAGR at 24% from 2021 to 2026[3]) sets demanding requirements on data communication technologies. New architectures and technologies will be also essential for the future development of 6G communications for improving the bandwidth and data transmission rate, while exhibiting lower latency and lower power consumption.

Optical interconnects enable higher bandwidth- distance products, higher bandwidth density, lower electromagnetic interference, and potentially lower power consumption than electrical interconnects. They are being deployed at increasingly shorter distances – for example, within and between data centres. In the longer term, chip-to-chip and even intra-chip communication may be performed with CMOS-compatible photonics. Beyond these applications, emerging precision applications – including atomic clocks, precision metrology, and transformative applications such as quantum communications and information processing – will also benefit from photonic capabilities integrated with electronics.

---

[1] https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks
[2] https://www.iea.org/commentaries/data-centres-and-energy-from-global-headlines-to-local-headaches
[3] https://blog.gitnux.com/internet-traffic-statistics/

These photonic capabilities range from silicon and heterogeneous III/V (membrane) photonics, to potentially disruptive technologies such as nanophotonics, and 2-D materials or graphene-based photonics. Additionally, developments in novel computing paradigms such as photonic and quantum computing threaten to make current security protocols insecure and will require the development of novel, future-proof cryptographic methods.

Other breakthroughs will concern adding intelligence close to the sensors (Intelligence at the edge) and/or to the data sources (IoT), and to integrate the components in a form factor that perfectly suits their applications. The initial generation of "Internet-of-Things" management was cloud-centric, where sensor data were collected from the periphery (or "edge"), then processed and analysed at the enterprise or platform tier. However, in that case, a tremendous amount of data needs to flow to the cloud and back, and a large amount of data processing power is required to structure and analyse it. In such a cloud-focused solution, latency and privacy concerns are often worrisome, or even prohibitive.

The term "embedded AI" or "edge AI" denotes how AI algorithms can be processed locally on a hardware device (e.g. a sensor) close to where the data is generated, and an action may then be required. A device using edge AI can process data it has collected and subsequently take decisions independently, without connecting to a central processing unit (CPU). Where initially local decisions will be supported by inference actions, there will be an evolution to training on the edge devices. Edge AI extends embedded computing, and contributes to economically effective solutions for the societal challenges we are facing in terms of:

- Reducing the energy consumption of the data infrastructure by transmitting only relevant data or pre-treated information (countering the unsustainable explosion of the energy demand by data centres and by telecommunication systems requiring higher bandwidths).
- Protecting personal data (GDPR compliance) by local processing and anonymisation of transmitted information.
- Increasing security and resilience due to a reduced reliance on telecommunication links as a result of local decision-making.
- Reducing latency by reducing the quantity of data needed to be transferred to and from a cloud, which is particularly important for automotive, digital society (real-time control of power distribution, for instance) and manufacturing applications, as well as some health applications.

Rethinking human activities to take advantage of the innovation opportunities offered by hyper-connectivity, AI solutions and new kinds of sensors based on miniaturised technologies will create numerous benefits for every new market, ranging from connected cars and digital health to smart home and smart living, and factories of the future. This should include lessons learned from the COVID-19 pandemic like the sudden increase in remote-working.

Sensors and biosensors will be an extensively studied discipline since their rapid, low-cost and highly sensitive features contribute to tremendous advances in many domains. Visible light, IR or multispectral imagers, lidar, radar and ultrasonic sensors, in combination with high-precision inertial sensors, will be essential for the deployment of advanced driver assistance systems (ADAS), augmented reality devices, and industrial automation for instance. Advancements in chemical-sensing technologies also open the door for multiple new markets. Gas sensors are increasingly integrated into IoT ecosystems to monitor air quality indoors and outdoors – for instance, wearable devices, smart city projects, sensor networks for pollution mapping, smart home electronics and automotive technology. Another key trend to utilise advanced gas-sensing technology is breath analysis, which aims at non-invasive diagnostics via detecting biomarkers from exhaled breath. Furthermore, pressure sensors in human and robot assisted minimally invasive surgery catheters are required to give haptic feedback to the surgeon. Miniaturized ultrasonic sensors open possibilities for minimally invasive medical imaging. However, to access the brain and smaller arteries in the body further miniaturization is required, posing a challenge for current pressure sensor technology.

R&D on highly selective biosensors will contribute to advances in next-generation healthcare, including personalised medicine and the ultrasensitive point-of-care detection of markers for diseases.

Next-generation electronic products are pushing the semiconductor industry to integrate more ultra-thin and flexible ICs. The combination of flexibility and processing capability is very desirable since it reduces weight and enables new form factors, while maintaining desirable functionality such as data logging and RF connectivity. Ultra-thin and flexible ICs enable more efficient and cost-effective solutions that will affect many applications, such as wireless communications, wearable electronics, implantable biomedical devices and the IoT.

The field of quantum sensing is rapidly expanding because quantum phenomena are extremely sensitive to their environment and thus can be used to measure physical properties with unprecedented precision. Quantum sensing refers to the process of employing an individual or an ensemble of quantum systems, often a quantum coherence and/or a quantum entanglement, to measure a physical quantity – ideally with improved accuracy, stability, sensitivity, sensitivity, precision, or spatial resolution compared to conventional measurements.

Quantum sensing usually describes one of the following:

1. Use of a quantum object to measure a physical quantity (classical or quantum). The quantum object is characterized by quantized discrete and resolvable energy levels. Specific examples include electronic, magnetic or vibrational states of superconducting or spin qubits, cold atoms, trapped ions, or photons.
2. Use of quantum coherence (i.e., wavelike spatial or temporal superposition of states) to measure a physical quantity.
3. Use of quantum entanglement to improve the sensitivity or precision of a measurement, beyond what is possible classically.

There are two generations of quantum sensors. The first, which includes devices such as microwave atomic clocks and superconducting quantum interference devices (SQUIDs), has been available for decades. The second generation, which includes atomic clocks, gravity sensors, magnetometers, gravimeters, gyroscopes, nitrogen-vacancy (NV) sensors, and other innovations, is just emerging. Second-generation quantum sensor applications may enable various domains:

- **Biomedical imaging**: neural sensing and heart imaging.
- **Spectroscopy**: imaging of molecular structure such as proteins.
- **Communications**: Signal receiving and amplification for radar communication; calibrating electrical standards to support 5G/6G.

- **Navigation**: Providing high-accuracy GPS; assisting with navigation inside buildings and underground.
- **Environmental monitoring**: Predicting volcanic disruption and measuring CO_2 emissions.
- **Infrastructure monitoring**: Monitoring mechanical stability and detecting leaks.
- **Geographical surveying**: Assisting with the location of oil and gas.

1.1.3 Major Challenges

To achieve application breakthroughs and strategic advantage, the European position must be reinforced through leadership in all relevant equipment, materials, processes and manufacturing technologies by driving the following Major Challenges:

- **Major Challenge 1:** Advanced computing, in-memory, neuromorphic, photonic and quantum computing concepts.
  Materials and substrates, process modules and integration technology for novel devices and circuits for advanced computing, memory and in-memory computing concepts based on nano-electronic, photonic, quantum or other technology.
- **Major Challenge 2:** Novel sensor, actuation and other devices that enable advanced functionality.
  Materials and substrates, process modules and wafer level integration technology for novel devices and circuits that enable advanced functionality: sensing including quantum sensing, actuating, power conversion, connectivity, etc.
- **Major Challenge 3:** Advanced integration solutions.
  Advanced integration including 2.5/3D integrated devices at wafer level, wafer-to-wafer (W2W) or material on wafer (sequential integration) or dies bonded/attached to a wafer (D2W), etc.
- **Major Challenge 4:** Advanced wafer fab equipment and manufacturing solutions.
  Equipment and manufacturing technologies for processing wafers in fabs from leading edge nodes to differentiated technologies and for advanced functionality devices, including new materials or unconventional geometry and heterogeneous integration technology options.
- **Major Challenge 5:** Advanced packaging, assembly & test equipment solutions.
  Equipment solutions to enable assembly and testing of a wide range of IC's from logic and memory to advanced 2.5/3D integrated devices.
- **Major Challenge 6:** Sustainable semiconductor manufacturing
  Solutions to reduce Greenhouse gas emission, water & gas consumption and use of hazardous materials.

1.1.3.1 Major Challenge 1: Advanced computing, in-memory, neuromorphic, photonic and quantum computing concepts

Semiconductor process technology and integration actions will focus on the introduction of new materials and substrates, process modules and integration technology, in close collaboration with the equipment, materials, modelling/simulation and embedded software communities, to allow for the necessary diversity in computing infrastructure. The applications range from high-performance cloud computing in servers for AI, ML and Gen AI data processing, edge computing/AI, and ultra-low power data processing at the IoT node level up to the highest possible performance. In order to lower the energy consumption of data transmission and processing, new currently investigated computing approaches such as in-memory and neuromorphic computing, or quantum computing (to solve classically unsolvable problems), the development of new materials, process modules and their integration in computing systems will be compulsory.

1.1.3.1.1 State of the art

The obvious solution for transistors with increased electrical performances is the use of fully depleted devices. The industry has adopted three integration methods: FDSOI CMOS, FinFET and GAA-style CMOS devices. In chip design it is now embraced that these transistor integration methods enable complementary roles depending on the level of system requirements, ie. cloud-based services, edge computing or extreme-edge device functionality.

FDSOI is a 2D technology based on a thin buried oxide (BOX) layer under the CMOS channel. FDSOI exhibits several advantages, such as reducing the leakage current at standby mode and its higher tolerance against soft errors compared to traditional structures. FDSOI is perfectly suited for ultra-low-power IoT automotive, edge AI and 5/6G SoCs. The leading companies currently produce 18 nm and 22nm FDSOI-based chips. The pilot line on FDSOI will develop 10 and 7nm FDSOI technology nodes to allow a further reduction in energy consumption, while increasing its information processing performances thanks to the integration of embedded non-volatile memories, RF modules and 3D-integration options.

The 3D based FinFET and GAA devices provide high current drive, and hence higher speed, low leakage and, most importantly, less wafer area per transistor than the classic 2D metal–oxide–semiconductor field effect transistor (MOSFET) technology. These new devices are designed and processed to deliver better performance for applications in high- growth markets such as hyper-scale data centres, AI and ML, autonomous vehicles and power-efficient SoCs for the most demanding computer applications. The international industry value chain is pushing production beyond the 2 nm node by moving towards more 3D-stacked FET architectures, and requires solutions in materials and process integration challenges (High-NA EUV, 3D-integration for instance) to realise these novel devices.

Scaling further has recently been demonstrated being possible via the introduction of CFET (complementary FET) or 3D-stacked FET devices. A future sub-2nm advanced System-on-Chip pilot line will enable bringing R&D of this technology to higher maturity including design enablement that will offer access to European manufacturing and design Industry. This allows SME's and start-ups to be competitive and strengthen EU's position in the value chain, increasing the resilience and maintain relevance in the world race of addressing the ever-increasing user needs and societal challenges.

A clear differentiation between logic, memory and process information in conventional von Neumann computing schemes necessitates the frequent movement of data between the memory and processor. Thus, much of the execution time and energy consumption is spent in the movement of data, a barrier referred to as the "von Neumann bottleneck", or "memory wall". This obstacle has been greatly exacerbated since the advent of data-intensive computing applications, such as ML and Gen-AI. Near-memory and in-memory computing are new paradigms, wherein the computing system is redesigned to process data at its storage – in the memory – thereby minimising the expensive movement of data.

Near-memory computing involves adding or integrating logic (e.g. accelerators, very small cores, reconfigurable logic) close to or inside the memory. Logic cores are usually placed inside the logic layer of 3D-stacked memories or at the memory controller.

Silicon and organic interposers allow separate logic chips to be placed in the same die package as a 3D-stacked memory while still taking advantage of the through-silicon via (TSV) bandwidth. Some foundries (Intel, Samsung, TSMC, etc.) offer this kind of heterogeneous integration.

Recent advances in heterogeneous integration technology focus on integrating Si-bridges to connect dies together over very short distances along the die perimeter (e.g. Intel EMIB technology, TSMC implementation of SoIC and "CoWOS" technology).

These heterogeneous integration technologies (Interposers, bridges, chiplets) are often referred to as 2.5D as they still place active die next to each other in a 2D plane. Through 3D integration technology stack active devices are placed vertically on top of each other. This is commonly done in the field of CMOS image sensors and HBM (High Bandwidth Memory) DRAM stacks. These 3D integration technologies use TSV technology in the active die, in combination with high-density die-to-wafer or wafer-to-wafer interconnect technologies. This greatly increases the number of functional interconnects per die area on the functional chips. 3D-Interconnect pitch below 5 µm and even down to 400 nm pitch have already been demonstrated.

An even denser 3D interconnect technology could be achieved by so-called "monolithic" 3D integration, a technique using sequential manufacturing of multiple layers of active devices and high density (Back-end of line) interconnect layers. Such monolithic integration is not currently available in foundries.

The amount of data processed in the cloud, the development of Internet-of-Things (IoT) applications, and growing data privacy concerns force the transition from cloud-based to edge-based processing. Limited energy and computational resources on edge push the transition from traditional von Neumann architectures to In-memory Computing (IMC), especially for machine learning and neural network applications. IMC also uses the intrinsic properties and operational principles of the memory cells and cell arrays, by inducing interactions between cells such that they can perform computations themselves. IMC aims to improve the energy efficiency of artificial/deep neural networks (ANN/DNN) hardware realizations by computing weighted-sum tasks in the memory arrays, for instance.

In an ANN chip, the neuron/synapse states are encoded as digital bits, clock cycles or voltage levels, while in a spiking neural network (SNN) chip, information is encoded into spike timing, to really mimic the biological brain operations. SNN hardware realizations are engineered to seek ultra-low power consumption and run at relatively low frequencies to emulate realistic biological behaviours, referred to as neuromorphic computing.

Due to the increasing need for large memory systems by modern applications (big data analytics, AI, etc.), dynamic random access memory (DRAM) and Flash memory scaling is being pushed to its practical limits. It is becoming more difficult to increase the density, reduce the latency and decrease the energy consumption of conventional DRAM and Flash memory architectures. 2D-NAND became monolithically integrated 3D-NAND, found the 3^{rd} dimension for scaling, and DRAM currently follows the same path. This will seek innovation in DRAM select transistor channel material (ALD MX_2, ALD Oxide Semiconductor etc.), capacitor dielectric, new cell architecture, new process steps etc. Alternative approaches are also being developed to overcome these barriers for implementing near- or in-memory and neuromorphic computing.

The first key approach consists of stacking multiple layers of memories (DRAM, Flash). With current manufacturing process technologies, thousands of TSVs can be placed within a single 3D-stacked memory chip. The TSV provide much greater internal memory bandwidth than the narrow memory channel. 3D-stacked DRAM and Flash are also commercially available.

The second major innovation is the use of emerging non-volatile memory (NVM) as parts of the main memory subsystem, and as embedded memories. New memory devices and technologies are currently being investigated that can both store data at high densities, lower costs, and present other benefits that will help computation with new in-memory computing paradigms: fast access time, long data retention, multilevel to analog ability and/or high endurance. The main emerging NVM technologies to allow in-memory computing architectures and embedded memories are: (i) phase-change memory (PCM), Threshold Change memory (TCM) ; (ii) magnetic RAM or spin-transfer or spin-orbit torque, or voltage-controlled magnetic anisotropy magnetic RAM (MRAM, STT-MRAM, SOT-MRAM, VCMA-MRAM); (iii) metal-oxide resistive RAM (RRAM or ReRAM) and conductive-bridge RAM (CBRAM) or memristors; (iv) ferroelectric FET (FeFET), RAM (FeRAM) and tunnel junctions (FTJ); (v) Electrochemical RAM (ECRAM) and (vi) Oxide Semiconductor (OSC) channel gain cell (2TnC, n=0 or 1). All these NVM types are expected to provide memory access latencies and energy usage that are competitive with - or close enough to - DRAM, while potentially enabling much larger non-volatility in main memory, and enabling new functionalities for computing systems outside Von Neuman architectures.

Thanks to its optimal compatibility to CMOS technology, Silicon Photonics is becoming a key material platform enabling technology for high-speed connectivity in data centres. In the near future, Photonic Integrated Circuits (PIC) can also bring significant changes to high-performance computers and unlocking the full potential of AI by resolving the transmission bottleneck of electronics.

Photonic processors that compute with photons instead of electrons display some extraordinary properties, such as an ultra-wide communication bandwidth, ultra-high processing frequency, and ultra-low energy consumption. Additional dimensions division multiplexing of light field such as wavelength and spatial mode enable multithread processing with almost no extra computing overhead, leading to a significant acceleration against traditional electronic computers.

By combining the high bandwidth and parallelism of photonic devices, photonic neural networks (PNNs) have the potential to be orders of magnitude faster than state-of-the-art electronic processors while consuming less energy per computation. PNNs aim to leverage high-speed optical devices to mimic essential computing primitives (neurons and synapses) and connect them into a neural network (ANN/SNN) with highly parallel and dense optical interconnects.

The goal of neuromorphic photonic processors is not to replace conventional computers, but to enable applications that require low latency, high bandwidth and low energies and are currently unreachable by conventional computing technology.

Photonic processors have light sources, passive and active devices. Nowadays, more than 10 companies produce PICs worldwide, mainly based on SOI wafers. However, there is no commercial fabrication platform that can simultaneously offer devices for light generation, optical amplification, wavelength multiplexing, photo-detection, and transistors on a single die. State-of-the-art devices in each of these categories use different photonic materials (SiN, Ge, InP, GaAs, 2D materials, etc) with incongruous fabrication processes (SOI, CMOS, FinFETs).

Energy efficient and fast switching optical and electro-optical materials are needed for non-volatile photonic storage and weighing (synapses), as well as high-speed optical switching and routing, with low power consumption. Neural non-linearities are already possible on mainstream platforms using electro-optic transfer functions, but new materials promise significant performance opportunities. Phase change materials (PCMs), graphene and ITO-based modulators can also be utilized for implementing non-linearities (neurons).

On-chip optical gain and power will require co-integration with active III-V (InP or GaAs) lasers and semiconductor optical amplifiers. Current approaches involve either III-V to silicon wafer bonding (heterogeneous integration) or co-packaging with precise assembly (see also Chapter 1.2). Quantum dot lasers are another promising approach as they can be grown directly onto silicon, but fabrication reliability does not currently reach commercial standards. Co-integrating CMOS controller chips with silicon photonics to provide electrical tuning control/stabilization and robust packaging for preventing PIC temperature fluctuations will be critical.

The past decade has seen an explosion in efforts to develop quantum computers that could revolutionize the fields of physics, medicine, biology, AI, finance and cryptography by exponentially speeding up certain computational domains. While such demonstrations certainly mark an essential technological milestone, tasks accomplished at such an unimaginable speed do not necessarily prelude the commercialization of quantum computers in the short term. To ensure the continued progression of quantum technologies over the next decade, advances in materials and fabrication processes are required for quantum computing hardware, following a path similar to the transistor technology scaling that enabled the evolution of digital computing.

1.1.3.1.2 Vision and expected outcome

Driven by market demand on the one hand for advanced high-performance computing, AI/ML, and on the other hand for mobility, edge AI and IoT devices, the advanced Si technology roadmaps for both FinFET/GAA and FDSOI will need to be pushed further. To enable this, a wealth of explorations into novel low-thermal-budget-processing 2D materials, nanowires, nanosheets or nanoribbons and quantum dots needs to be combined with significant developments in advanced 3D integration schemes of materials and devices. In parallel, to overcome the von Neumann bottleneck, development of new computing paradigms such as neuromorphic, in-memory, photonic and quantum computing is essential.

New memory concepts will support the correct memory hierarchy in various applications. An example here is the opportunity to push new memory concepts (resistive RAM (RRAM), phase-change RAM (PCRAM), Thresholds Change RAM, STT- MRAM, FeFET, FeRAM, FTJ, Electro-Chemical RAM (EC-RAM)) to the demonstration level in the IoT infrastructure (from server, over edge to nodes). These alternative memories require the development of advanced novel materials (magnetic, phase-change, nanofilament, ferroelectric, electrochemical). A much closer collaboration between new material innovation, process, device integration, device teams and system architects is indispensable in the future. New markets will require storage class memory to bridge the performance gap between DRAM and NAND Flash. Edge AI and IoT applications will require low-power embedded devices and cloud computing with more mass-storage space. The standard memory hierarchy is challenged. Indeed learning algorithm requirement will dramatically increase the requirements on memories densities, access bandwidth and endurance on one hand, while new functionality such as Multilevel to analog non-volatile memory cells are required to allow an efficient in memory computing approach like ANN, from direct matrix vector multiplication to CNN or spike neural architectures. Simultaneously, advanced interconnect, System on Chip (SoC) integration issues will need to be addressed (cf. also Major challenges 2 and 3), with innovative solutions to reduce costs being required. The option to use advanced 3D and optical input/output (I/O) technological solutions to circumvent limitations of traditional I/O architectures are strengths to foster and build upon in Europe.

Furthermore, the downscaling of photonic components (light sources, modulators, photodetectors, optical phase shifters…) and their integration on CMOS platforms would lead to photonic neural networks (PNNs) for overcoming the current memory bottleneck. Such PNNs are expected to achieve orders of magnitude enhancement in energy efficiency and throughput compared to ANNs and SNNs. Hence, the intimate collaboration between the strong EU electronic and photonic communities would enhance the EU leadership in this domain.

One of the current major issues for most of the developed quantum technologies is their future scalability. For example, quantum computing error rates multiply as scale increases, and creating large numbers of qubits that are stable enough for long enough is extremely challenging. Collaboration between quantum "laboratories" and the EU electronic industry looks compulsory for strengthening the EU force in quantum computing.

To maintain the European competencies in advanced design for integrated circuits and systems, a close link with a strong effort in semiconductor process technology and integration has to be maintained. Issues such as the creation of standards for the IoT, reliability for safety or mission-critical applications, security and privacy requirements need close collaboration among all players to build leadership going forward in this coming generation of advanced and distributed computing infrastructure and diversified system performance.

Expected achievements

Maintaining competence on advanced logic and memory technology in Europe is key to maintaining strategic autonomy and supporting societal benefits from the core technology base. Implementation of dedicated and sustainable pilot lines for specialised logic processes and devices supporting European critical applications is also a major objective, as is the exploration of new devices and architectures for low-power or harsh environment applications.

1.1.3.1.3 Key focus areas

This challenge includes the following key focus areas:

Topic 1.1

- Explorations of the scaled Si technology roadmaps of the 2 and 1 nm nodes including FinFET/Trigate and stacked gate-all-around horizontal or vertical nanowires, Forksheet-, complementary FET architectures, next generations FDSOI, 3D integration, and further device and pitch scaling where parallel conduction paths (nanowires, nanosheets, nanoribbons, etc.) are brought even closer together. It includes novel device-interconnect technology such as - but not limited to - contact from wafer backside.

Topic 1.2

- Exploration and implementation of materials beyond Si (SiGe, SiC, GaN, Ge, InGaAs, InP, functional oxides, 2D material heterostructures, CNT and nanowires).
- Unconventional devices and materials, such as 2D and III-V materials, oxide-semiconductors, metamaterials, metasurfaces, nanowires, CNTs, nanosheets, nanoribbons, nanoparticles, quantum dots, spin effects, functional oxides, ferroelectric and magnetic, which are being investigated to overcome the limits of conventional CMOS logic and memories.

Topic 1.3

- Novel device, circuit and systems concepts for optimum PPAC specifications, high-energy efficiency and novel paradigms such as for near/in-memory, neuromorphic, optical and quantum computing.
- New Photonic integrated circuits (PIC) for a further improvement of optical I/Os for enabling compute, memory, and networking ASICs to communicate with dramatically increased bandwidth, lower latency and at a fraction of the power of existing electrical I/O solutions.
- In a longer term, Photonic Neural Networks (PNNs) could bring the benefits of photonics over electronics (energy efficiency, low latency, higher bandwidth, parallelism…) to AI/ML.

Topic 1.4

- Long-term challenges such as steep slope switches (tunnel FET, negative capacitance FET, nanoelectromechanical systems / NEMS), spin-based transistors, and alternative high-performance switches.

Topic 1.5

- New embedded non-volatile memory (eNVM) technologies to enable local AI processing and storage of configuration data, which decrease data transmission volume, energy needs and allows for more efficient control of electric powertrains and batteries, along with many other applications in the IoT and secure devices domains.

1.1.3.2 Major Challenge 2: Novel sensor, actuation and other devices that enable advanced functionality

Materials and substrates, process modules and wafer level integration technology for novel devices and circuits that enable advanced functionality: sensing - including quantum sensing, actuating, power conversion, connectivity, etc.

1.1.3.2.1 State of the art

Besides the highly integrated chips necessary to overcome Major Challenge 1 on advanced computing, memory and in-memory computing concepts, many more devices are needed to achieve advanced functionalities – such as sensing and actuating, power management, and interfaces to other systems. This is what has also been named “More-than-Moore” for the past 20 years, and is an integral part of all systems, as well as one of the strengths of European microelectronics. Given the inherently diverse nature of this sector, the state of the art will be captured by providing a few snapshots of key technologies.

For IoT applications, logic and RF functions are combined, but not with the highest efficiency required by the ultra-long lifetime of unattended objects. Energy harvesting schemes, often based on photovoltaics, do exist, yet are not always able to provide the requested energy supplement of self-contained low-volume and low-cost sensor nodes.

Smart optical, mechanical and magnetic sensors are already able to provide a wealth of information for complex systems. Nevertheless, there are current limits to integrating various types of sensors monolithically. In the field of optical sensors, for instance, depth mapping requires complex scanning schemes using either mechanical systems or large volume and poorly integrated light sources. Devices based on rare or expensive materials, which are not compatible with standard CMOS technology, cover various useful zones of the electromagnetic spectrum. The same is true for chemical-sensing technologies, which are mostly based on metal oxides or other coated materials. While solutions for specific gases and applications are starting to emerge, sensitive and robust technologies using semiconductors still remain to be developed for a large number of applications and species, or concentration ranges, selectivity and stability. The situation is similar for many kinds of sensors and actuators. For instance, fine pitch displays are beginning to be possible, but will require new advances both in high brightness low variation sources and assembly methods.

Technologies used for the next generation of quantum sensing include neutral atoms, trapped ions, spin qubits, superconducting devices, and photonics. Please note that in the previous sentence “quantum sensing” means sensing using individual quantum objects and their entanglement. They can be used to measure the magnetic field (B), the electric field (E), the rotation (θ), the acceleration (g), the time (t), the frequency (f), the temperature (T), the pressure (P), the force (F), the mass (m) and the voltage (V) (Table 1). Charged systems, like trapped ions, will be sensitive to electrical fields, while spin-based systems will mainly respond to magnetic fields. Some quantum sensors may respond to several physical parameters.

In the more conventional light sensing techniques, sensors are often based on the absorption of a given semiconductor species, which leads to a rather broad but limited energy range. Emerging techniques to tailor the wavelength range by either confinement of the sensing element (“quantum dots”) or the combination of different materials in the same stack. The question of wavelength selection, in case a further wavelength selectivity is needed, for instance to address both the visible and near infrared bands, is usually addressed by stacking different sensing elements and optical filters of various principles.

In power technologies, recent years have seen the emergence of wide bandgap materials able to reduce the losses of power conversion, namely SiC and GaN, or other wide bandgap technologies. These technologies are making quick inroads as one of the cornerstones for the energy transition, and are becoming dominant in some sectors, like high-power electric motors... However, they are still nascent, and the challenge is to develop low-cost (involving larger diameter, good quality and less-expensive substrates) and robust technologies. Today, SiC is produced mainly on 150 mm substrates, while some GaN devices are produced on silicon substrates, but the technology and epitaxy techniques will still need further refinement (and even breakthroughs). Moreover, the development of disruptive substrate technologies as well as layer transfer will be key steps toward a cost-effective, high-performance solution linked with transition to 200 mm then from 200 mm to 300 mm substrates where possible and where volumes justify it. This is essential for future integrated logic and power management functions using technologies combining logic and power transistors. Besides research on wide bandgap materials, the Si-based insulated-gate bipolar transistor (IGBT) technologies have further innovation potential in the area of cost-sensitive applications. Challenges are in the domain of high-power and high-voltage electronics with high junction temperatures processed toward larger diameter substrates – leading to increased power densities and lower costs to support the transformation in the energy systems with Si-based power semiconductors. In some cases the power dissipation of the devices - which remains high despite their better efficiency - is a real challenge and needs to be addressed both at device and wafer level and at the packaging level (see also Chapter 1.2).

For RF and communication technologies, recent advances in integrating RF technologies on low-loss substrates such as SOI have allowed the integration of switches as well as amplifiers on the same silicon substrates. This concept is in production in Europe on 200 mm and 300 mm wafer substrates. Further advances are on the way on both substrate diameters, which will allow the integration of more functions and address the requirements of complex 5G systems below and beyond 6 GHz, up to the mm waveband. Synthetic antennae systems for radar or communications are emerging thanks to highly integrated RF technologies, including BiCMOS, but are often limited by power consumption and costs. For example, BiCMOS brings advantages over some III-V technologies for some applications, leveraging the capability to integrate digital functions. For sure, it cannot compete with GaN when very high power is requested but on the other hand GaN cannot compete with BiCMOS for the products requiring mixed-signal functions. New, RF technologies delivering high output power and efficiency like RF GaN (GaN-SiC as well as GaN-Si) and other III-V based materials - such as InP - could overcome these limitations. Those technologies have to be considered either stand-alone or in combination with either RFCMOS or BiCMOS using heterogenous integration. In the field of communications, the integration of photonics technologies with electronics is gaining commercial ground. Further advances in efficient source integration, and modulation and power efficiency, are still needed to use them more widely. New advances in fine photon handling can also open the way to innovative sensing techniques, either by using light interaction with the signal to be measured outside the PIC and sensing the effect of this interaction within the PIC or by using the PIC as an actuator, for instance by implementing optical (phased array) antennae.

In optical communications, the intrinsic capability of light waves to transmit signals with low latency and power dissipation, at ultra-high data rates, can be scaled from long-haul infrastructures to intra-datacenter optical links, down to chip-to-chip photonic interconnects. However, bulk silicon cannot meet the necessary requirements of these integrated optics applications which can be addressed with silicon-on-insulator (SOI) technology. SOI photonics, which is part of Silicon Photonics, is one route to enable the development of novel lidar systems as well as support advancements in quantum technologies. By leveraging mature semiconductor manufacturing methods, engineered wafers that incorporate SOI technology offer a powerful approach toward broader adoption of advanced chip-scale integrated optics. In addition to substrate progress a host of innovations will be necessary and are expected to boost the (Silicon) Photonics Integrated Circuits (PIC). Among those are new generations of active devices including the heterogeneous integration of active devices (sources, modulators) or passive components and detectors (mainly for sensing), for example for wavelengths outside the communication frequencies. They involve new materials and integration techniques.

In "traditional" polyimide (PI)-based flexible electronics, the continuing trend is towards more complex designs and large-area processing, especially in displays and sensor arrays. Since the achievement of high-performing flexible electronics by monolithic approaches is limited, hybrid approaches are used when conventional electronics (such as thinned chips) is assembled on flexible electronic substrates. For more complex devices, the reliability and performance of organic materials or mechanical and processing properties of inorganic materials are still a focus of research activities in addition to adapted and optimised assembly techniques. In general, current R&D activities indicate that technical spots can be identified where a merging of novel flexible devices and adapted Si electronics create progress beyond the state of the art.

1.1.3.2.2 Vision and expected outcome

Depending on the application, the advantages of More-than-Moore (or SoC) technology are size, performance, cost, reliability, security and simpler logistics. Therefore, this approach is a key enabler for European industry. To maintain and strengthen Europe's position, it is necessary to improve existing technologies, and to seamlessly integrate emerging technologies in a reliable and competitive way. All application domains addressed by the ECS SRIA will benefit from components with very diverse functionalities.

Specific process technology platforms may be required, as in the case of biomedical devices for minimally invasive healthcare or point-of-care diagnosis, or mission-critical devices in automotive, avionics and space. Semiconductor process and integration technologies for enabling heterogeneous SoC functionality will focus on the introduction of advanced functional (nano-)materials providing additional functionalities and advanced device concepts.

Innovations for these domains require the exploration and functional integration, preferably in CMOS-compatible processing, of novel materials. A non-exhaustive materials list includes wide bandgap materials, III-V semiconductor compounds, 2D materials (e.g. graphene, MoS_2 and other transition metal dichalcogenides), 1D (e.g. nanowires, carbon nanotubes) and 0D (e.g. nanoparticles, quantum dots) materials, metal oxides, organic semiconductors, ferro- and piezoelectric, thermoelectric and magnetic thin films materials, materials with specific properties (for optics: dielectrics, bi-refringent, phase change materials which can also be used for conductivity change properties, for electromagnetic properties: low permittivity, etc..), metamaterials and metasurfaces. Obviously, safety and environmental aspects should also be taken into consideration.

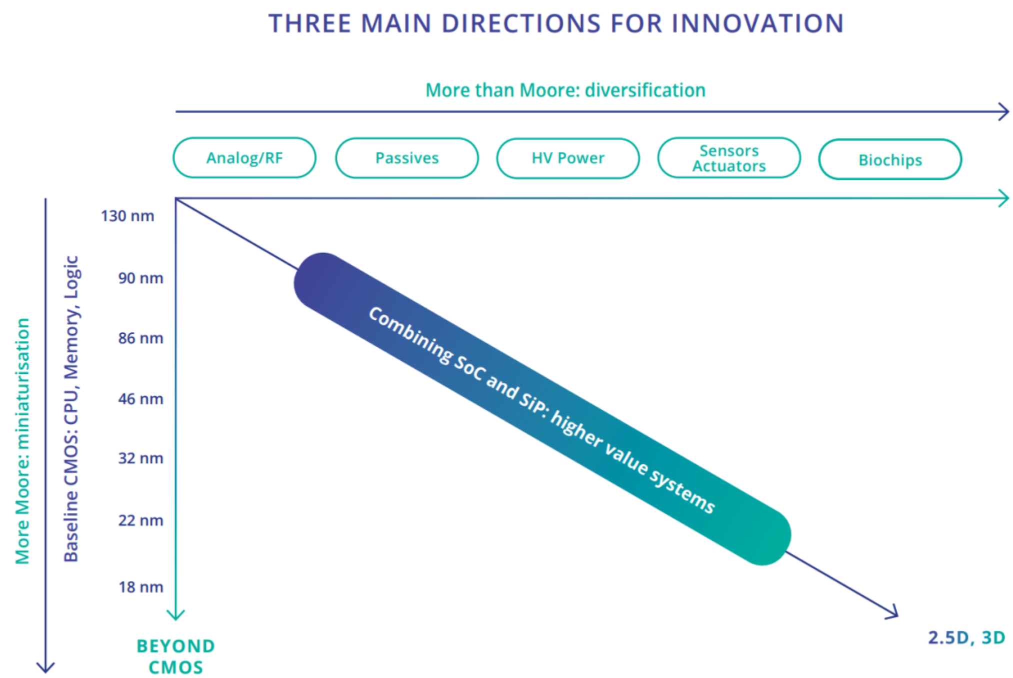


Figure 1.1.2 Diversification of applications, continued miniaturisation and integration on chips and in package leads to higher value systems (Source: ST Microelectronics/IRDS)

The driver for SoC integration is always a clear demand from the application domain. To maintain and push forward Europe's position, the focus should be on emerging technologies as they are introduced, as well as new developments in the equipment and materials industry, in which Europe has a leading position. Furthermore, the early generation of models and their initial validation for benchmarking and intellectual property (IP) generation are required to reinforce position of Europe in specific design concepts and architecture, especially when used in combination with re-use IP and third-party IP blocks to secure fast time to market.

Expected achievements

This will involve the implementation of pilot lines for integrated application-defined sensors, novel IoT solutions, complex sensor systems and new (bio)medical devices, new RF and mm-wave device options (including radar and sensing), much more efficient power management, higher integration of passive devices, photonics circuits and options, novel and more efficient displays and electronics, this list being non exhaustive. Key will be the initiation of process technology platforms for the exploration and exploitation of advanced functionalities through integration of novel reliable materials.

The exploration and implementation of materials beyond Si will require strategic collaborative EU projects for European industry to become more independent, and will result in the development of an EU-based supply chain for wide-bandgap materials, for example, including a move towards larger substrate sizes of 200 and 300 mm (i.e. SiC and GaN, and also InP which provides a further challenge to GaN in that it cannot be grown on Si).

1.1.3.2.3 Key focus areas

More specifically, the following challenges are identified (this is a non-exhaustive list).

Topic 2.1: Application-specific logic: as explicitly treated in sub-section 1.1.4.3, heterogeneous SoC integration can require specific solutions for logic to be integrated with More-than-Moore technologies such as the following:

- Logic integration with RF, optical or sensor technologies.
- Integration of lasers and detectors within silicon photonics platform.

Topic 2.2: Advanced sensor and actuators technologies:

- Mechanical sensors and actuators (e.g. acceleration, gyroscopes, microphones and microspeakers).
- Chemical sensor devices such as selective gas-sensing components for environmental monitoring or smart medicine and smart health (e.g. CO, CO_2, NO_x, O_3, toluene, VOCs, acetone, H_2S, etc…).
- Physical sensors (magnetic, optical, RF).
- Multispectral or highly sensitive optical sensors.
- Transmitter/receiver technologies for applications such as lidar and active phased array imaging.
- Biomedical and biochemical sensors.
- New, more efficient displays, in particular micro displays.
- environmental protection technologies for audio MEMS and integrated technologies for enhanced robustness against outer environment for MEMS based audio devices.

Topic 2.3: Advanced power electronics technologies (Si-based, BCD, SiC, GaN, Ga_2O_3, AlN etc.) to enhance the efficiency of motors, energy storage, lighting systems, etc. More specifically:

- Higher power density and frequency, wide-bandgap materials for high temperature electronics, new CMOS/IGBT processes, integrated logic, uni- and bipolar; high voltage classes, lateral to vertical architectures.
- Materials for energy harvesting (e.g. perovskite solar cells, piezoelectric ceramics and thin films) and storage (e.g. perovskites, ferroelectrics and relaxors), micro-batteries, supercapacitors and wireless power transfer.
- Power devices and modules for highly demanding automotive, industrial and energy infrastructure applications.
- Substrates towards larger diameters to serve future greater demand for cost-sensitive power solutions.

Topic 2.4: Quantum sensor technologies:

- Atomic vapor or cold atom (Bose-Einstein condensate) based sensors.
- Trapped ion-based quantum sensors.
- Solid state spin-based quantum sensors making use of Nitrogen-Vacancies or color centres as sensitive element.
- Superconducting circuit sensors based on SQUIDS, flux qubits and charge qubits.
- Sensors based on photon entanglement using nonlinear optical media.

Topic 2.5: Advanced RF and photonics communication technologies to interface between semiconductors components, subsystems and systems. These technologies should enable better and more energy-efficient control of emission and reception channels (for example, for 5G connectivity and 6G) via:

- New energy-efficient RF and mm-wave integrated device options, including radar (building on e.g. SiGe/BiCMOS, FD SOI, CMOS, GaN or other III-V compounds, PIC).
- Development and characterisation of new RF cryogenic electronics for Quantum Information Processing (QIP), as well as logic devices at quantum-enabling cryogenic temperatures, taking into account the available cooling power of refrigerators and interfacing requirements at different operation temperatures.
- Energy-efficient computing and communication, including a focus on developing new technologies.
- Bringing MOEMS and micro-optics, nanophotonics, optical interconnections, photonics-enabled device and system options into a CMOS-compatible manufacturing flow.

1.1.3.3 Major Challenge 3: Advanced integration solutions

Advanced integration including 2.5/3D integrated devices at wafer level, wafer-to-wafer (W2W) or material on wafer (sequential integration) or dies bonded/attached to a wafer (D2W), etc.

By splitting the chip into smaller functional physically separate parts, the overall system yield improves and system performance is enhanced. In addition, by using system-independent IP block design and verification, as well as common die-to-die interfaces (including IP re-use and use of third-party IP), a faster time to market can be achieved. One counterpart being that the design and testing strategy needs to be entirely revised in order not to lose the yield advantage, and that assembly costs (as collective as possible) and yields must also be mastered. This is the key overarching challenge of this approach.

1.1.3.3.1 State of the art

Over the last few years a huge variety of semiconductor products have emerged where several functions are added in one IC, enabled by advances in integration technology. This is the so-called System In Package (SiP) approach.

To maximise the benefits from ICs made of small geometry nodes, below 40 nm typically and certainly as of 7 nm and less, there has already been a move toward more advanced methods, to manage complexity in the most cost-effective way. These advanced integration methods involve technologies such as flip chip, but also - depending on the use cases - wafer-level packaging, fan-out wafer-level packages without substrate interposers and complex 3D structures with TSVs, micro-bumps and thin dice.

The functional diversification of technologies, where digital electronics meets areas such as analog, photonic and MEMS technologies, has been advanced through the assembly of heterogeneous elements. For example, in today's power stages in automotive powertrain applications, power modules integrate several dice in parallel. Similarly, 5G networks are enabled by advanced RF functionality, often combining a photonic interface with in-package integrated logic and memory functionalities. Semiconductor materials encapsuled in packaging technology have already moved from being largely silicon-based or based on III-V compounds for photonic or high-power RF applications, to more advanced SiC and GaN compounds. On the part of the package, the industry has moved towards environmentally friendly lead (Pb) and halogen-free moulding compounds. For wire bonding, a similar move from aluminium and gold towards copper and silver wiring has been made. Furthermore, flip chip attach has made a transition to lead-free bumps (inside the package) and BGA using lead-free balls (at the interface between the package and the board) materials.

1.1.3.3.2 Vision and expected outcome

This challenge covers the integration of new chip technologies in advanced low parasitic packages, as well as chips of different functionalities resulting from the previous two challenges – e.g. CMOS logic, NVM, NEMS/ MEMS, RF, analogue, sensing, actuating, optical, power management, energy harvesting and storage – into a SiP.

Depending on the application and type of system, the key drivers and parameters to be improved can be the density of contacts, the parasitics, the integration of passives - including antennae -, the thermal dissipation, the optical quality, the number of dice to be integrated either horizontally or vertically, the ability to handle various environmental conditions including extreme temperatures or resistance to chemicals, etc. ..It is becoming more and more clear to everyone that the overall performance of the system is dependent not only on the semiconductor technologies but also - and sometimes in equal part or even predominantly - on the integration technology.

Advanced integration technologies are required for mm-wave applications (> 30 GHz), both GaN/Si RF and other high-electron-mobility-transistor (HEMT) devices, or dedicated MEMS and sensor devices (e.g. electro-optics for lidar without moving parts) and display technologies used in AR/VR. Depending on the application, heterogeneous integration technology can provide a better compromise between available functions, performance, cost and time to market.

System integration technologies are a key enabler for European industry, including for instance the essential field of energy transition with power management devices, but also including longer term evolutions like for instance the new field of cryogenic QIP, characterisation of logic devices at quantum-enabling cryogenic temperatures, and associated packaging challenges. To maintain and strengthen Europe’s position, it is necessary to improve existing technologies and develop emerging technologies, as well as to integrate both to advanced electronic systems in a competitive and reliable way. All application domains addressed by the ECS agenda will benefit from innovative integration technologies.

Moreover, component carriers also known as Integrated Circuit (IC) Substrates represent a big portion of the package cost (up to 50%), excluding the semiconductor component itself. This is particularly true for Flip Chip Ball Grid Array (FC-BGA) Substrates, which are essential component carriers in high-performance, high-interconnect density computing solutions. Due to the increase of functionality and driven by chiplets integration, FC-BGA substrates are facing many challenges not only in terms of miniaturization but they are also becoming central elements for thermal management in advanced computing systems.

Integration of the above functionalities in miniaturised packages and (sub)SiP require fundamental insights into application needs and system architecture. Process and characterisation technology to realise this integration is part of this third Major Challenge, and is essential for ensuring Europe’s prominent role in supplying novel solutions for the various existing and emerging application domains.

At the macro-scale level, a system consists of a collection of large functional blocks. These blocks have quite different performance requirements (analogue, high voltage, eNVM, advanced CMOS, fast static

RAM (SRAM), multi-sensing capability, etc) and technology roadmaps. Therefore, for many applications it is of increasing interest to split the system into heterogeneous parts, each realised by optimum technologies at lower cost per function, and assembled with parts using high-density 3D interconnect processes. In other words, for each application or system the SoC / SiP trade-offs and partitioning have to be revisited in the light of the respective evolutions of base technologies.

It is clear that 3D integration in electronic systems can be realised at different levels of the interconnect hierarchy, each having a different vertical interconnect density. Different technologies are therefore required at different levels of this 3D hierarchy.

For the reasons above the Electronic Design Automation of 3D integration is to be much further explored and constitutes a very challenging domain as described in Chapter 2.3.

1.1.3.3.3 Key focus areas

Research and development priorities are focused on innovative approaches, such as the following.

Topic 3.1: Advanced Si interconnect technologies:

- Interconnect technologies that allow vertical as well as horizontal integration: this includes process technologies for vertical interconnects, such as TSV, through-encapsulant via (TEV), through-glass via technologies and microbumps, and copper/copper bonding, as well as process technologies for horizontal interconnects such as thin film technologies for redistribution on chips.
- Implementation of advanced nanomaterials and metamaterials, including low-thermal-budget-processing 2D materials, nanowires, nanoparticles and quantum dots with scalable logic and memory device technologies.
- New materials to maximize efficiency of the Package Thermal dissipation (Flip Chip Thermal Enhanced BGA) is required to deal with the very demanding computing applications in extreme conditions i.e. Automotive, space environment.
- New materials like Al bumps from Electroplating could solve reliability issues and ensure advanced CMOS compatible integration.
- Power electronic substrates will benefit from thick Al layers fabricated by novel electroplating technologies, and other thermal redistribution techniques.

Topic 3.2: Specific sensing, actuation and display technologies.

- Solutions for advanced optical functionalities on wafer, either for sensors (visible, NIR, infrared), for PIC, or for AR/VR, near-eye and head up displays, or any optical system.
- Specific solutions for the signal conditioning of other sensors (mechanical, physical, chemical…)
- Increasing functionality in IC Substrates for high efficiency power delivery including voltage regulator circuitry and integrated capacitances and other passives.

Topic 3.3: 3D integration technologies:

- High-integration density and performance-driven 3D integration (power/speed). For this category, denser 3D integration technologies are required: from the chip I/O-pad level 3D-SIP, to finer grain partitioning of the 3D-SOC and the ultimate transistor-level 3D-IC (see Sub-section 2.3.1 for the 3D landscape).

- Chip-package-board co-design. This will be of utmost importance for introducing innovative products efficiently with a short time to market (and which is closely linked to the work described in Section 2.2).
- System integration partitioning: the choice of 3D interconnect level(s) has a significant impact on system design and the required 3D technology, resulting in a strong interaction need between system design and technology with a significant impact on electronic design automation (EDA) tools.

System requirements and semiconductor device technology (Major Challenges 1 and 2) will evolve at the same time, creating momentum for further interconnect pitch scaling for 3D integration technology platforms. Hence, the timelines of all four challenges of this section are strongly connected.

1.1.3.4 Major Challenge 4: Advanced wafer fab equipment solutions

This is about equipment and manufacturing technologies for processing wafers in fabs from leading edge nodes to differentiated or mature technologies and for advanced functionality devices, including new materials or unconventional geometry and heterogeneous integration technology options.

The semiconductor equipment and manufacturing sector in Europe provides the global market with best- in-class equipment to enable the manufacturing of miniaturised electronic components. The European equipment industry, RTOs and small and medium-sized enterprises (SMEs) active in this sector have a long history of successful mechanical engineering, tailor-made machinery, optical equipment, metrology, inspection and testing equipment, and chemical processing tools. This history of success is prominent in several domains, foremost in lithography (in particular EUV) and metrology, but also in thermal processing, deposition, etching, cleaning and wafer handling.

1.1.3.4.1 State of the art

At the forefront of semiconductor manufacturing equipment is the production of logic and high-performance memory, which are applied mainly in portable devices as well as advanced cloud computing and data storage facilities. The continuous increase of device density known as Moore's law is being driven by an ability to create ever-smaller features on wafers. The technology leaps required to keep up with Moore's law have already been achieved via additional roadmaps complementing ongoing 2D pattern size reductions. They are realised by combining various devices, materials, and 3D and system architecture aspects that required dedicated long-term investment in high-tech equipment solutions. Enabled by current deposition, lithography, etch, processing and metrology tools and their performance, the 3nm technology node is in production by market leaders, solutions for 2nm node are planned for 2025[4], and even Angstrom nodes[5] are being explored.

For the production of miniaturised and reliable More-than-Moore electronic components and systems, such as sensors and sensor systems, MEMS, advanced imagers, power electronics devices, automotive electronics, embedded memory devices, mm-wave technologies, and advanced low-power RF technology, many equipment and manufacturing solutions have been implemented.

---

[4] TSMC 2nm Update: N2 In 2025 (anandtech.com) 3nm Technology - Taiwan Semiconductor Manufacturing Company Limited (tsmc.com)

[5] Intel Demonstrates Breakthroughs in Next-Generation Transistor... Intel Accelerated

1.1.3.4.2 Vision and expected outcome

The ever-increasing demand for leading-edge logic and memory technology is driving the development of new equipment and material solutions for sub-2 nm node semiconductor technologies. Besides finding equipment solutions for further shrinking minimum feature sizes well below 1 nm, the alignment accuracy of successive layers, called “overlay” or in another definition type “Edge Placement Error”, needs to move towards Angstrom levels in a process technology roadmap that combines complex materials in 3D structures and architectures. At the same time, productivity demands on the equipment continue to increase to maintain reduced overall production costs. Process yield also continues to be a challenge with shrinking feature size and the increasing impact of defects and contamination.

In the realm of cutting-edge technology, the advent into new nodes necessitates the development of novel materials. These materials are either integrated into the structure to enhance electrical properties or utilized as sacrificial elements to facilitate the creation of intricate 3D shapes. The precision in material deposition is becoming increasingly critical, requiring meticulous control over where materials are applied, their thickness, and whether they are deposited on horizontal or vertical surfaces.

The semiconductor industry is driven by the dual goals of boosting productivity and slashing energy and water consumption. Wafer fabs are striving to process more wafers per hour without increasing power usage, achieved through enhanced equipment throughput, real-time optimization of production flows and streamlined processes. Despite variations in throughput across different equipment types—such as lithography’s DUV throughput at approximately 295 wafers per hour whereas throughput in deposition might take up to a few minutes for single wafer processes—there is a universal push for higher throughput and yield, alongside refined process control to minimize additional steps and reduce defective wafers. This pursuit has led to the development of more sophisticated equipment in patterning, processing and metrology, with kWh/wafer serving as a pivotal metric for sustainable innovation.

The increasing technical complexity calls for more efficient design and production methods. At factory level, solutions based on advanced mathematics, statistics or AI should enable extension across current domains enabling a wider scale integration. Advanced data modelling and analysis techniques, combined with generative AI should rapidly enable smarter detection of deviations (whether its process or product quality, equipment or global production performance) together with faster reaction thanks to AI augmented troubleshooting and guided analysis.

The intricate nature of future semiconductor manufacturing equipment requires a substantial increase in skilled labor for design and realization. Without advancements in efficiency, the local labor market may fall short in providing the necessary expertise, potentially slowing down innovation and affecting profit margins. There also generative AI should be instrumental in mining existing knowledge deposits to support transfer to new generations. New methods will have to be developed to ensure the efficiency of this transfer and the sustainability of European know-how. While the use of generative AI should also enable the integration of this knowledge into software assistants or copilots and eventually autonomous systems, techniques will have to be developed to ensure that the system will stay in (human) control.

Automation trends like Industry 4.0 and 5.0 are reshaping manufacturing, demanding greater data exchange and secure integration of design data into production processes. This evolution aligns with the Smart Industry roadmap.

Environmental considerations are also at the forefront, with efforts to eliminate toxic substances such as PFAS-containing materials and SF6 gas from equipment and processes. Companies are committed to ambitious NetZero goals, focusing on enhancing energy, water, and material efficiency while exploring recycling opportunities for gases like hydrogen and carbon dioxide, as well as processed materials and consumables. Collaborations with equipment and materials suppliers are key to these initiatives, which include refurbishing and reusing components.

Overall, these developments pose significant challenges to the global semiconductor industry across various domains but also present opportunities for European equipment vendors and their supplier networks.

The overarching goal of equipment development is to lead the world in miniaturisation techniques by providing appropriate products two years ahead of the shrink roadmap of the world's leading semiconductor device and components manufacturers[6]. Internationally developed roadmaps such as the International Roadmap for Devices and Systems (IRDS) will also be taken into consideration[7]. Currently, leading integrated device manufacturers (IDMs) are forecasting a continuation of the technology roadmap following Moore's law at least until 2029[8], which corresponds to at least four new generations after the current technology node.

All-in-all, this represents a major challenge to the international semiconductor industry in the areas of lithography, material innovation, processing, assembly, process control, analysis and testing, as well as an opportunity for the well-positioned European equipment vendors and their network of suppliers.

1.1.3.4.3 Key focus areas

The key focus areas for innovative semiconductor manufacturing equipment technologies are as follows.

Topic 4.1 Wafer fabrication equipment

- Advanced patterning equipment for sub-2 nm node wafer processing using deep ultraviolet (DUV) and EUV lithography, and corresponding subsystems and infrastructure (e.g. pellicles, masks and resist).
- Mask manufacturing equipment for sub-2 nm node mask patterning and tuning, defect inspection and repair, metrology and cleaning.
- Advanced holistic lithography using DUV, EUV and next-generation lithography techniques, such as e-beam and mask-less lithography, directed self-assembly (DSA) and nano-imprinting.
- In-line wafer inspection and analysis for process control, defect detection, contaminant control etc.
- Multi-dimensional metrology (MDM) and inspection for sub-2 nm node devices that combine all the spectrum of physical tools and data processing techniques.

---

6 https://investor.tsmc.com/english/encrypt/files/encrypt_file/qr/phase5_support/TSMC%201Q20%20transcript.pdf

7 https://irds.ieee.org/

8 https://www.anandtech.com/show/15217/ intels-manufacturing-roadmap-from-2019- to-2029

- Thin film processes including thin film and atomic layer deposition, doping and material modification, and corresponding equipment and materials, able to support the increase of binary, ternary and quaternary materials
- “Bottom up” technologies to selectively deposit materials on topography or on a selected material.
- Integrated surface preparation, deposition and etch process technologies for optimal interface engineering.
- Innovative equipment and process strategies for perfect gapfill of metals and dielectrics in decreasing feature sizes.
- Equipment and manufacturing technology for wet and dry processing, wet and dry etching, including (atomic layer) selective etch processing, thermal treatment, laser annealing and wafer preparation.
- Increased utilization of AI and modelling (e.g. computational chemistry) techniques for material and process development.

Topic 4.2: Wafer fab equipment for differentiated and/or mature technologies

- Technologies and tools for the manufacturing and integration of semiconductor components made with advanced nanomaterials and metamaterials (low-thermal-budget-processing 2D materials, nanowires, nanoparticles, quantum dots, etc) with logic and memory technologies.
- High-volume manufacturing tools for the production of III-V, GaN, SiC or other material substrates of up to 200 mm, or 300 mm in the future.
- Enable productivity enhancements (e.g. wafer diameter conversions) for heterogeneous integration technologies to significantly improve cost-competitiveness.
- New manufacturing techniques combining chip and packaging technologies (e.g. chip embedding), which will also require new manufacturing logistics and technologies (e.g. panel moulding).
- Dedicated equipment for manufacturing of electronics on flexible, structural and/or bio-compatible substrates.
- New electroplating equipment for ionic liquids (i.e. for Al deposition).

Topic 4.3: Automated manufacturing technologies including fab robotics and wafer transport & handling are required to enable IC-fabs with interconnected tools to support flexible, sustainable, agile and competitive high-volume semiconductor manufacturing[9]

- Enable flexible line management for high-mix and distributed manufacturing lines, including lines for fabrication and deposition of advanced functional (nano) materials.
- Develop infrastructure technology to interconnect systems from various companies… And anticipate, configure, optimize, etc.
- Implement fast (and deep) learning as well as semi-automated AI-based decision- making to control processes, to enhance quality, increase reliability, shorten time to stable yield, preserve knowledge and master complexity.

[9] 2023 IRDS Factory Integration (ieee.org)

- Apply Productivity Aware Design (PAD) approaches with a focus on predictive maintenance, virtual metrology, factory simulation and scheduling, wafer-handling automation and the digitisation of the value chain for AI-based decision management.
- Doubling semiconductor manufacturing in Europe in 2030 also means evolving and upgrading installed base through incremental approaches, which will necessarily mean increased complexity. Managing such hybrid factories will require advanced decision support and diagnosis techniques leveraging GenAI, Retrieval Augmented Generation (RAG) and integrating existing human knowledge and know how.
- Developing comprehensive modelling and sharing techniques, to enable seamless flow and utilization of information across the whole value chain, will require significant evolution of the existing knowledge management techniques and technologies (NLP to exploit existing documentation, diffusion and sharing of cutting-edge or strategic knowledge).

Topic 4.4: Sustainable semiconductor manufacturing

- Future innovations should also address new environmentally friendly solutions for manufacturing (e.g. in terms of energy consumption, chemical usage) and environmentally friendly new materials (e.g. in terms of quality, functionality, defects) in parallel with addressing the continued cost of ownership challenges. This will entail, for example, new precursors, chemicals for deposition and other wafer-processing materials, as well as gas delivery, gas handling, pumps and abatement systems. This will also comprise the study and implementation of new solutions for both effluent segregation as described above, including PFAs, as well as for recycling of water and other chemicals.

Major Challenge 5: Advanced packaging, assembly & test equipment solutions

Semiconductor packaging and assembly equipment refers to the machinery and tools used in the process of packaging and assembling integrated circuits (ICs) and other microelectronic devices. Packaging is a crucial part of semiconductor manufacturing as it protects the semiconductor die, connects the chip to a board or other chips, and conducts heat dissipated by the components it contains. It affects power, performance, and cost on both macro and micro levels. The equipment includes various types of machinery from wafer dicing, encapsulation (or moulding) of the die into a plastic package, wire-bonding (on thermal compression, ultrasonic, and stitch bonding), wireless bonding (quasi-monolithic, hybrid bonding) to package marking and labelling equipment.

Semiconductor test equipment is used to evaluate the functionality and performance of semiconductor devices, such as integrated circuits (ICs), during and after the manufacturing process. This equipment is essential for ensuring that the devices meet the required specifications and quality standards before they are shipped to customers. The main types of semiconductor test equipment include Automatic Test Equipment (ATE), Wafer Test Equipment, In-Circuit Test (ICT) etc. These tools are crucial for detecting defects, ensuring reliability, and maintaining high yield rates in semiconductor manufacturing.

1.1.3.4.4 State of the art

The current state of the art in semiconductor packaging and assembly equipment is highly advanced and includes several key technologies that have emerged over the past two decades. Current advanced packaging technology uses sophisticated methods to aggregate components from various wafers, creating a single electronic device with superior performance. It is particularly important for applications like 5G, autonomous vehicles, IoT technologies, and virtual/augmented reality, which require high-performance, low-power chips capable of processing large amounts of data. [10] 3D Integration: Techniques such as 2.5D, 3D, fan-out, and system on-a-chip (SoC) packaging have been developed to supplement traditional wire-bonding and flip-chip technologies. These methods allow for higher density and performance by stacking chips vertically or integrating multiple chips into a single package. In the specific case of chiplets, the approach involves designing chips in a modular fashion that can be integrated with other types of chips to create a complete system. It allows for more flexibility in design and can lead to improvements in performance and power efficiency.

Yet, while reducing the complexity of design, 3D integration - when done at wide scale - should considerably complexify production and supply chain management. Integrating multiple chips in a single package will introduce a new level of complexity in assembly lines which will have to synchronize multiple flows, both internally (bring the right quantity and quality to the machine) and externally, i.e., at supply chain level, across various front-end facilities or even companies. Designing and operating robust supply chains will require digital twins operating at much wider scale.

The current state of the art in semiconductor test equipment is characterized by several key advancements. Machine Learning has been integrated into test systems to analyze large volumes of test data in real-time. This allows for the detection of subtle defects or performance deviations that might be missed by traditional methods. Similarly, Machine Learning algorithms predict when automated test equipment (ATE) is likely to fail, enabling proactive maintenance to minimize downtime. With the rollout of 5G networks and the proliferation of IoT devices, test equipment providers have developed innovative solutions to meet the demands for higher data rates, lower power consumption, and greater reliability and availability.

These advancements in packaging and test equipment have helped ensure that semiconductor devices meet the necessary quality and performance standards. This is critical for the reliability and functionality of modern electronics. As the industry continues to evolve, we can expect further innovations in semiconductor packaging, assembly and test equipment to enhance performance, reduce costs, and improve reliability.

1.1.3.4.5 Vision and expected outcome

The future developments for semiconductor packaging and test equipment are expected to focus on several key areas which may reshape the industry. [11]

As devices become smaller and more complex, advanced packaging technologies like system in-package (SiP) and 3D stacking are becoming important. The advanced multichip packaging approach integrates multiple semiconductor components into a single package, which is well-suited for applications such as mobile devices, automotive computing, and generative artificial intelligence (GenAI). It aims to improve performance and time-to-market while reducing manufacturing costs and power consumption. High-Density Fan-Out, 2.5D IC, and 3D IC Packaging Technologies are becoming increasingly crucial for future data-centric applications. To maximize the benefits from ICs fabricated in nodes of 7 nm and below, the

---

[10] https://www.mckinsey.com/industries/semiconductors/our-insights/advanced-chip-packaging-how-manufacturers-can-play-to-win
[11] https://www.bcg.com/publications/2024/advanced-packaging-is-reshaping-the-chip-industry

equipment will need to enable complex 3D structures with Through Silicon Vertical Interconnect Accesses ("TSVIA"), micro-bumps and thin dies, as well as wafer-to-wafer bonding, to speed up production of 3D ICs.

Packaging equipment must be modified to meet the decreasing feature sizes and increasing precision requirements of advanced packaging. This may lead to introduction of packaging equipment in the front-end part of semiconductor manufacturing and, hence, increased requirements to contamination control.

Further equipment developments include functional diversification of technologies, where digital electronics meet the analog world, using advanced assembly/packaging of heterogeneous pieces of chips ("chiplets") and of chips, sensors and/or smart antenna components. It should be noted that these trends are closely related to the trends formulated for (front-end) wafer-fab equipment, amongst others creation of silicon vias and backside power distribution networks.

An ongoing focus on enhancing functionalities of heterogenous systems by combining "traditional" semiconductors with iphotonics components or chiplets and addressing high speed component transfer by the assembly industry, with ambitious goals of developing transfer capability up to 1 million components per hour (without physical manipulation) by 2028.

Besides this technological frontier, other trends for Assembly equipment will also play a role. For example:

- Factory automation to increase product reliability with multi-facetted device inspections, sorting and advanced tracking and tracing as well as data storage throughout the whole of the production lines.
- Product developments for highly specialized applications such as power modules, medical applications, organ on chip devices, communication and agrifood.
- Development and application of 3D printing technology in heterogeneous integration, covering materials, process and equipment.

The future developments for semiconductor test & inspection equipment are anticipated to be influenced by several emerging trends and technological advancements. Building on the existing state of the art, further integration of AI and Machine Learning will play a significant role in transforming test engineering and the connected equipment. They will enable real-time analysis of extensive test data, identify subtle defects, and optimize test patterns to reduce test durations and enhance efficiency. As devices become smaller and more complex, new testing strategies and specialized test fixtures will be developed to access intricate device connections and ensure quality and performance in advanced packaging technologies. Test equipment designers will also need to explore and integrate innovative approaches to enable improved wafer-level testing & inspection to catch defects early in the manufacturing process, which is crucial for maintaining high standards of quality.

These advancements by EU high-tech equipment industry will help ensure that semiconductor devices meet the necessary quality and performance standards, which is critical for the reliability and functionality of modern electronics.

1.1.3.4.6 Key focus areas

Topic 5.1: Packaging & assembly equipment:

- Equipment and manufacturing technology supporting 3D integration and interconnect capabilities such as chip-to-wafer stacking, fan-out WLP, multi-die packaging, "2.5D" interposers, wafer-to-wafer sequential processing, TSVs and transistor stacking.
- Enhanced equipment optimised for high-volume manufacturing of large batches of the same package into efficient reconfigurable equipment for the manufacturing of different packages in smaller batches.
- New process tools for die separation, attachment, thinning, handling and encapsulation for reliable heterogeneous integration on chip and in package, as well as assembly and packaging of electronics on flexible substrates.
- Equipment development to suit the requirements of multi-component assembly on flexible and stretchable substrates, especially in roll-to-roll for both conductive adhesives and soldering.
- New selective bonding equipment based on inductive or reactive heating.
- Equipment solutions to handle glass substrates and enable Through Glass Via's and glass plating.

Topic 5.2: Test and inspection equipment:

- In-line and off-line technologies for the Testing, Validation and Verification (TV&V) of heterogeneous chips and SiP with ever-increasing number of features and ever-decreasing feature size to tackle the challenge of failure localisation in these highly complex (packaged) chips.
- Characterisation and inspection equipment for quality control at multiple levels and different scales of semiconductor structures, films and components.

1.1.1.1 Major Challenge 6: Sustainability of semiconductor manufacturing

The semiconductor industry is living a paradox. On the one hand, continuous advances in chip capabilities are propelling the global effort to reduce carbon emissions through electrification and energy efficiency improvements in devices and all types of equipment. On the other hand, semiconductor manufacturing causes significant emissions itself, responsible for large CO_2 and Green House Gas (GHG) outputs. As an example, the carbon footprint of ICT was evaluated at about 1200–2200 $MtCO_2$ eq. in 2020, or equivalently 2.1%–3.9% of the world greenhouse gas (GHG) emissions. A significant part of this footprint, i.e., between 360 and 660 $MtCO_2$ eq., is attributed to the semiconductor-manufacturing phase[12].

At the same time, the semiconductor manufacturing industry must answer to the unprecedented demand for semiconductors across the industry for big emerging technologies such as AI/ML, Gen. AI, IoT, EVs and 5G/6G networks. This demand may double in the next four to six years. Therefore, reducing the CO_2 and GHG emission in semiconductor manufacturing is of utmost importance for the semiconductor industry.

In addition or coupled to CO_2 and GHG emission, semiconductor manufacturing is resource intensive in energy, water, chemicals and raw materials. Moreover, the use of energy, water and other resources by

[12] T. Pirson et al. 2023, "The Environmental Footprint of IC Production: Review, Analysis, and Lessons From Historical Trends", IEEE Trans. on Semicon. Manuf. 36, 1, 56-67

the semiconductor industry continues to increase as wafer processing becomes more complex and as the industry expands production worldwide.

Furthermore, the semiconductor industry uses PFAs-containing materials in numerous critical applications. They are primarily utilized in the photolithography process to facilitate the smooth transfer of patterns onto the semiconductor substrate. PFAs stands for per- and poly-fluoroalkyl substances, known as “forever chemicals” (PFAs have lifetimes in the thousands of years), and refers to synthetic chemical compounds that contain multiple fluorine atoms attached to an alkyl chain. The broad definition of PFAs by the Organization of Economic Cooperation and Development encompasses more than 10,000 unique chemicals. Because of human health and environmental factors associated with the persistence, bioaccumulation and toxicity of several PFAs, legislative and regulatory efforts worldwide are seeking to initiate restrictions that could limit the use of PFAs-containing materials.

1.1.1.1.1 State of the art

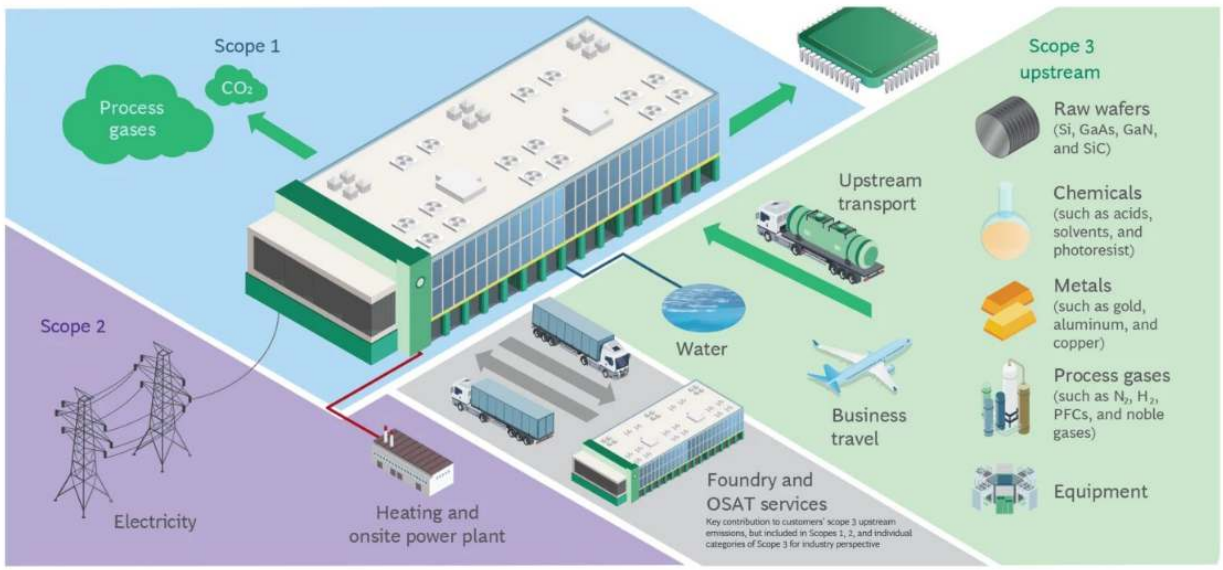
For decades, the ECS community spent many efforts for enhancing the energy efficiency of electronic components for computing and sensing thanks to aggressive improvements to circuit design, component integration, and software, as well as power-management schemes and integration of specialized accelerators. However, the flip side is that higher energy efficiency does not necessarily translate to lower CO_2 and GHG emissions from semiconductor manufacturing processes.

According to the GHG Protocol, there are three types of GHG emissions (Fig. 1.1.3):

- Scope 1 – direct emissions from business operations and manufacturing.
- Scope 2 – indirect emissions including those from power plants to meet business energy requirements
- Scope 3 – emissions including all other indirect emissions in a company’s value chain; upstream emissions are those generated by suppliers or their products, while downstream emissions are related to the usage of products containing semiconductors during their lifetime. Scope 3 downstream emission is nearly independent of the semiconductor-manufacturing phase.

In a typical semiconductor fab, 35% of GHG emissions fall into the Scope 1 category, compared with 45% for Scope 2, and 20% for Scope 3 upstream[13].

[13] “Keeping the semiconductor industry on the path to net zero”, McKinsey & Company November 2022



Source: BCG analysis.

Note: Si = silicon; GaAs = gallium arsenide; GaN = gallium nitride; SiC = silicon carbide; N_2 = nitrogen; H_2 = hydrogen; PFCs = perfluorinated compounds; OSAT = outsourced semiconductor assembly and test.

Figure 1.1.3 GHG emission sources from the semiconductor manufacturing

With about 80% of semiconductor manufacturing emissions falling into either scope 1 or scope 2 categories, semiconductor manufacturing lines control a large portion of their GHG emission profile. 85-90% of scope 1 emissions arise from fluorinated process gases used during wafer etching and cleaning, and chamber cleaning, while 10-15% originate from the use of fluorinated heat-transfer fluids. These gases, which include perfluorocarbons (PFCs), hydrofluorocarbons (HFCs), NF_3, SF_6 and N_20 have high global-warming potential (GWP) as well as a long atmospheric lifetime. Moreover, their emission rises as node size shrinks. According to the sixth assessment report of the Intergovernmental Panel on Climate Change, GWP100 values (defined as a relative GWP value for the period of 100 years with respect to CO_2) of CF_4, C_4F_8, NF_3 and SF_6 are 6630, 10200, 16100 and 23500, respectively.

To reduce PFC emissions, several strategies, such as alternatives, process optimization, recycling/recovery and abatement have been attempted.

The use of alternative materials to PFCs is expected to reduce PFC emissions by employing environmentally benign chemistries with low GWPs. However, finding relevant alternative materials will require significant R&D efforts. Process optimization through increasing the efficiency of PFC-consuming processes by adjusting process parameters, such as temperature and chamber pressure, can reduce PFC usage and, correspondingly, PFC emissions. Recycling/recovery and abatement of PFCs from exhaust streams can also control PFC emissions.

Scope 2 emissions, which represent the highest proportion of GHG from semiconductor companies, are linked to the energy required to run their extensive production facilities. The sources of these emissions include the following:

- Tool fleets containing hundreds of manufacturing tools, such as lithography equipment, ion implanters, and high-temperature furnaces. Their energy consumption represents the most part (39%) of a typical semiconductor fab[14].
- Large clean rooms requiring climate and humidity control with overpressure and particle filtration.
- Extensive sub-fab facilities for gas abatement, exhaust pumps, water chillers, and water purification.

As the node size of chips continues to shrink, energy requirements at production facilities are expected to rise significantly. Indeed, the energy consumption for manufacturing of a 28nm and a 3nm 300mm wafer has been modelled to be 400 and 1400 kWh, respectively[15]. The facility energy was not included in this model.

Lowering greenhouse gas emissions is one of the semiconductor industry's most complicated challenges due to the energy needed to make chips and the unique requirements of the manufacturing processes. This complexity is even enhanced as the semiconductor industry must also expand globally to meet the growing demand for chips.

Semiconductor manufacturing is also extremely-water intensive and requires large volumes of ultra-pure water to avoid the contamination of electronic devices. A typical semiconductor manufacturing facility uses 7.5 to 15 million litres of ultra-pure water per day, and it takes about 6,000 litres of city water to create 3,800 litres of ultrapure water. In response to prolonged droughts and increasing demand for semiconductors, the semiconductor industry is focusing on new ways to recycle, reduce, and reuse the water used in their production. New advancements in water treatment have emerged to allow semiconductor manufacturers to recover and reuse wastewater and remove targeted contaminants. They allow reaching a water recycling rate, defined as the ratio between the recycled water and the total water use (which is the sum of recycled water use and water intake), ranging from 12 to 70%[16]. However, since the demand for semiconductors will strongly grow in the coming years, the semiconductor industry must improve recycling water rate, especially in water-stressed countries.

Given its carbon-fluorine chemistry, PFAs-containing materials offer a unique set of surface tension, stability and chemical compatibility that many semiconductor applications require. The following paragraphs highlight key uses and challenges of PFAs in each of the application groupings, which are critical to the semiconductor industry:

- **Photolithography** materials containing PFAs are a critical component used within photoacid generators in chemically amplified resists and bottom antireflective coatings, top antireflective coatings, surfactants, barrier layers, photo-imageable polybenzoxazoles and polyimides for dielectric and buffer coat applications, and photoresist applications. The amount of PFAs

---

[14] S-C. Hu et al. 2020, "Energy savings approaches for high-tech manufacturing factories", Case Studies in Thermal Engineering 17, 100569

[15] M. Garcia Bardon et al., "DTCO including Sustainability: Power-Performance-Area-Cost-Environmental score (PPACE) Analysis for Logic Technologies," 2020 IEDM, San Francisco, CA, USA, 2020,

[16] Wang Q. et al. 2023, "Water strategies and practices for sustainable development in the semiconductor industry", Water Cycle 4, 12–16

consumed in each technology node depends on the number of masks used and the complexity of the multi-patterning steps.

- **Wet chemistries** are applied within semiconductor manufacturing for cleaning, stripping, wet etching, chemical mechanical planarization, metal plating, and to facilitate other processes. Most wet chemistry applications do not contain PFAS. However, there are some applications that rely on PFAs materials for specific performance requirements. Examples of these applications include - but are not limited to - post-plasma photoresist strip, high aspect ratio collapse mitigation, selective film inhibition, wetting of low surface energy substrates, and specific parts cleaning.
- **Dry etching and deposition process**: Perfluorocarbons (PFCs) and hydrofluorocarbons (HFCs) are essential gases for directional etching and cleaning of silicon compounds. Additionally, fluorinated organometallic compounds are essential for the deposition of metal-containing films. Some PFCs and HFCs are part of the PFAs category such as CF_4, C_2F_6, C_4F_8 and C_3F_8.
- **Heat Transfer Fluids** (HTFs): Many semiconductor-manufacturing processes entail physical and chemical processes that require precisely controlled temperatures, and thus are highly reliant on HTFs. In both cooling and heating applications, fluorinated HTFs (F-HTFs) help ensure the ability to provide the precise temperature control required in specific manufacturing operations within the semiconductor fabrication process.
- **Semiconductor manufacturing and related equipment** and infrastructure articles. Semiconductor manufacturing facilities and the manufacturing equipment and infrastructure within them contain a multitude of articles. An article is any object made from one or more substances and mixtures which during production is given a special shape, surface or design that determines its function to a greater degree than its chemical composition, whether on its own or in an assembly with other articles, substances and mixtures. PFAs-containing articles include those made of a fluoropolymer, articles coated or painted with a fluoropolymer, or other PFAs-containing materials (such as oligomers) and those made of non-PFAs polymers containing PFAs processing/machining aids or additives. Many semiconductor manufacturing applications require the use of PFAs-containing articles for safety, contamination control, resilience and other factors.
- **Pump fluids and lubricants**. Semiconductor manufacturing relies on the extensive use of robotics, automation and vacuum systems to achieve nanometre-scale precision. The use of lubricants, many of which need to be fluorinated, is essential to the precision and reliability of these systems.

Due to the extremely wide range of individual PFAs and their respective use cases in the semiconductor manufacturing process, time frames for finding and deploying alternatives can be very long, even in best-case scenarios. Depending on the use case, an alternative that has the same functionalities must first be invented and tools and/or process facilities must be adapted before said alternative could be qualified and deployed. As an example, some alternatives are currently developed for substituting PFAs in HTFs, in working fluids for vacuum pumps and in some articles in equipment. However, no alternatives to PFAs are available in photoresists, developers, antireflective coatings, rinsing solutions, wet and dry etching[17].

The short-term approach to reduce or even prevent the emission of PFAs from the semiconductor industry is to remove PFAS-containing materials from fab water, solvent wastes, gases and chemical vapours.

---

[17] Lay D. et al., “A guide to PFAS in electronics”, ChemSec H2020 Project No. SA22171340

1.1.1.1.2 Vision and expected outcome

Since GHG emissions from a typical semiconductor fab arise from its power/energy consumption, identifying strategies for reducing power consumption in a semiconductor manufacturing line looks mandatory. Process optimization is one approach to reduce the energy consumption. It requires a close collaboration between chip manufacturers, R&D labs and tool suppliers by simultaneously optimizing yield and energy consumption during the least energy-efficient process steps. Similarly, such collaborations should also help equipment suppliers to manufacture more energy-efficient equipment and materials.

As the energy and water consumptions of a typical semiconductor manufacturing line depend on the number of process steps, and evolve proportionally[18], reducing the energy consumption through process optimization should also lower the water consumption. However, the most convenient approach to lower the water intake of the semiconductor industry, is to enhance the current water recycling rate in order to reach a rate higher than 70% in the coming years.

Finding alternative gases for replacing PFCs will be long and costly, because only few green solutions are now viable alternatives to current gases. It seems to be easier to find alternatives for fluorinated gases for cleaning processes (pre-clean, seasoning, chamber cleans) than for dry etching. Looking for such viable alternatives will require narrow collaboration along the semiconductor supply chain from R&D labs and chemical suppliers to the semiconductor manufacturers.

Hence, over the short to mid-term, the current solution to decrease GHG emissions from fluorinated gas usage is the implementation of abatement systems to contain and decompose the problematic compounds. The current abatement efficiency in a semiconductor manufacturing line typically ranges from 20 to 65%[19]. This abatement rate should be enhanced in a short term. This will remain the case until alternative gases with fewer emissions are available, or until gas recycling is widely adopted.

Gas recycling is another approach for reducing GHG emission from the semiconductor industry. It consists in capturing unutilized process gases and by-products through various means, such as membrane separation, cryogenic recovery, adsorption, and desorption. They can then refine them into pure process gases that can be used again, potentially reducing process-gas emissions. For this lever to become economically viable, researchers will need to address major challenges related to the separation of process-gas outflows and purification.

Concerning PFAs, the semiconductor industry recognizes the need for and is undertaking additional R&D to:

- Characterize the human health and environmental risks associated with PFAS-containing materials used in the industry.
- Develop analytical methods to characterize PFAS-containing materials.
- Evaluate PFAS releases to air and/or water.

---

[18] Wang Q. et al. 2023, "Environmental data and facts in the semiconductor manufacturing industry: An unexpected high water and energy consumption situation", Water Cycle 4, 47–54

[19] S. Jones "IEDM 2023 – Modeling 300mm Wafer Fab Carbon Emissions", https://semiwiki.com/semiconductor-services/techinsights/340325-iedm-2023-modeling-300mm-wafer-fab-carbon-emissions

- Identify, test and implement substitutes to either eliminate PFAS-containing materials, or substitute PFAS-containing materials with those having lower human health or environmental risks.
- Evaluate and test abatement technologies to capture or destroy PFAS-containing materials before their release to the environment.

1.1.1.1.3 Key focus areas

Topic 6.1: Semiconductor manufacturing line:

- Enhanced water recycling rate including the full removal of PFAs from wastewater.
- Enhanced efficiency of gas abatement systems for reducing GHG and PFAs emissions in the atmosphere.
- Optimized process steps for reducing the energy consumption of the least energy-efficient steps, while maintaining high manufacturing yield.

Topic 6.2: Equipment and material suppliers:

- Improve energy efficiency of the equipment.
- Implement re-use of parts/modules for equipment maintenance and servicing.
- Reduce the use of PFAs-based materials in equipment production.
- Find viable alternatives to PFAs for etching and layer deposition.
- Find alternatives to PFAs for photolithography resists and wet chemistries.

1.1.4 Timeline

All leading European industry and research actors should align their activities with international roadmaps and timelines. Roadmap exercises are being conducted in various projects and communities, including NEREID[20] and the IEEE's IRDS[21], in which European academia, RTOs and industry are participating. For system integration, the International Electronics Manufacturing Initiative (iNEMI)[22] and the new Heterogeneous Integration Roadmap activities are also considered. The European R&D priorities are planned in synchronisation with global timeframes and developments that are under continuous adaptation. The timelines below are high-level derivatives from these global evolutions, and follow the structure of the four Major Challenges described above.

For Major Challenge 1, the roadmap for process technology and device/system integration presents relatively clear timelines, although economic factors will determine the speed of adoption in industrial manufacturing. Dedicated process technologies (e.g. low-power and high-operating temperature) will follow feature scaling with some delay, focusing on other performance indicators. Areas where the roadmaps and timelines are less clear (e.g. new computing paradigms) will be introduced at low technology readiness levels (TRLs).

---

[20] https://www.nereid-h2020.eu/
[21] https://irds.ieee.org/
[22] https://www.inemi.org/ https://eps.ieee.org/technology/ heterogeneous-integration-roadmap/2019-edition.html

For Major Challenges 2 and 3, the timeline of the implementation of new technologies largely depends on the needs and roadmaps of the systems, and will result from the interaction within application-driven projects and test-bed initiatives. The timing of new equipment and manufacturing solutions for these challenges should be derived from the schedules of the major European semiconductor manufacturers. This includes roadmaps for key future semiconductor domains, such as automotive, healthcare, safety and security, power, MEMS, image sensors, biochips, organ-on-a-chip, photonics, lighting, etc. Fast implementation and modification of these new device technologies will pave the way for the technologies of tomorrow.

First, the development of sub-2nm solutions in terms of equipment and materials as part of Major Challenge 4 needs to be two-to-three years ahead of mass adoption, and is of critical importance to maintaining European leadership. Second, new equipment and materials solutions should be developed in line with the needs defined in the roadmaps of Major Challenges 1–3. Lastly, improving manufacturing efficiency and enhancing yield and reliability are ongoing tasks that need to be performed in accordance with the needs of the “more-Moore” and “more-than-Moore” domains. Fundamentals of “manufacturing science” will concern projects at rather low TRLs (typically 3–5), whereas implementation in pilot lines and full-scale production lines will contemplate higher TRL projects (typically 7–8). For most of the manufacturing science projects, the execution will take place in the medium- to long-term timespan, although shorter-term impact, such as improving the uptime of equipment due to Productivity Aware Design or the improvement of robustness of the manufacturing processes, will get due attention to enhance competitiveness.

Major challenge	Topic	Short term (2025 – 2029)	Mid term (2030-2034)	Long term (2035 and beyond)
Major challenge 1: Advanced computing, in-memory, neuromorphic, photonic and quantum computing concepts	**Topic 1.1:** Extensions of the scaled Si technology roadmaps High-performance Ultra-low power 3D integration	Sub-N2 R&D 2nd generation gate-all-around devices, forksheet integration 18 nm FDSOI at technology platform integration level	N1,5 R&D - 3rd generation of Gate-All-Around devices CFET introduction 12/10 nm FDSOI at technology platform integration level 3D monolithic integration	Sub-1 nm node logic and memory technology (nanowires, nanosheets) at process and device research level Vertically stacked nanosheets 3D monolithic integration Beyond 10 nm FDSOI at technology platform integration level
	Topic 1.2: Exploration and implementation of unconventional devices based on materials beyond Si	SiGe (high Ge) channel Cu alternative solutions	Ge channel Optical interconnects 2D materials exploration	III-V channel Low-thermal-budget-processing 2D materials device integration
	Topic 1.3: Novel device, circuit and systems concepts, such as for near/in-memory, neuromorphic, optical and quantum computing neuromorphic, optical and quantum computing	Near/in-memory computing 3D heterogeneous integration (logic/memory)	In-memory computing Neuromorphic computing (spiking) 3D monolithic integration Photonic SOI	Quantum computing Optical computing
	Topic 1.4: Long-term challenges such as steep-slope switches, spin-based transistors and alternatives		TFET CNTFET 2D material FET	NCFET NEMS switch Topologic insulator electronic devices Spin wave devices Mott FET (VO2, HfO2, etc)
	Topic 1.5: New eNVM technologies	PCRAM STT-MRAM FDSOI embedded MRAM and/or PCRAM	PCRAM PCRAM (MLC) SOT-MRAM ReRAM FeRAM	ReRAM (MLC) Hi-density ReRAM TCRAM VCMA-MRAM ECRAM
Major challenge 2: Novel sensor, actuation and other devices that enable advanced functionality	**Topic 2.1:** Application-specific logic integration	ULP 18 nm FDSOI technology with embedded memory integration	12/10 nm FDSOI technology integration with memories New technology integration architectures for neuromorphic computing 3D monolithic or quasi monolithic stacking integration for heterogeneous elements	3D monolithic integration
	Topic 2.2: Advanced sensor and actuator technologies	Continuous improvement of sensitivity (imagers, IMU, etc), range (lidar), and reduction of individual sensor area and energy consumption Development of miniaturised low power chemical sensors Development of biomedical sensors integrated with micro/ nanofluidics Heterogeneous and quasi monolithic integration of sensor technologies with (ULP) logic/memory technologies, including 28/22 nm FDSOI for the IoT environmental protection technologies for audio MEMS	Quantum sensors (see 2.4) Ultra-low power chemical sensor systems for pollution monitoring Energy autonomous connected sensor systems Integrated biomedical sensor systems Heterogeneous and quasi monolithic integration of sensor technologies with novel device, circuit and systems memory and computing concepts, including 12nm FD-SOI for IoT Integrated technologies for enhanced robustness against environmental effects for MEMS based audio devices	Nanoelectronic sensor devices with individual molecule sensitivity and selectivity Nanoelectronic biomedical sensor systems Monolithic integration of sensor technologies with novel devices, circuit and systems memory and computing concepts 10 or Beyond 10 nm FDSOI for IoT

	Topic 2.3: Advanced power electronics technologies	Silicon, BCD, SiC and GaN-based technologies and substrate materials Energy-efficient systems, including energy harvesting	New CMOS and IGBT processes Smart GaN devices (combining logic and power devices) Vertical GaN power devices Towards 300 mm GaN and 200 mm SiC substrates Energy-autonomous systems Energy harvesting and energy storage systems	B-Ga_2O_3, AlN Diamond
	Topic 2.4: Quantum sensing technologies		Integration of cells used in trapped ions, neutral and cold atoms technologies, and entangled photons in a CMOS platform coupled to laser sources and (single) photon detectors Superconductor qubits Maintaining performance when manufactured at chip-scale and optimising SWaP-C (size, weight, power and cost).	Development of entangled qubits for improving the sensitivity growth of "large" diamond single crystals investigation of color centers-based qubits in wide bandgap semiconductors such as SiC, GaN
	Topic 2.5: Advanced RF and photonics communication technologies	Enable 5G connectivity RF and mm-wave integrated device options building on, for example, SiGe/BiCMOS (increase of ft), RF and FDSOI, CMOS, PIC,GaN/Si, GaN/SOI and GaN/SiC technologies Towards 300 mm GaN-Si substrates Next-generation SOI for mm-wave photonics SOI 200 mm POI	Improve RF front-End components roadmap (switches, LNA, antenna tuners) Strained materials Tiny silicon thickness and uniformity Improve linearity substrate behaviour RF substrate options for advanced CMOS nodes Integration of III-V semiconductors on silicon and SOI Integration of III-V semiconductors on photonics SOI RF interposer for heterogeneous integration combining III-V and CMOS 3D stacking of different functions (RF with digital, ...) New materials for advanced functions	
Major challenge 3: Advanced integration and solutions	**Topic 3.1:** Advanced interconnect technologies	Vertical as well as horizontal integration via TSV, TEV, microbumps Fan-out WLP or embedded wafer-level BGAs and chip- embedding in laminate materials Advanced wafer-stacking technologies Packaging & bonding technologies with advanced thermal management capability	3D stacking/horizontal connecting of dies/chiplets Advanced nanomaterials (including low-thermal-budget-processing 2D materials, nanowires, nanoparticles, etc) Critical raw materials elimination from packaging bill of materials such as W, Co, Mo, Be, BeO	Co-Packaged Optics, Active photonic interposers
	Topic 3.2: Specific power, sensor and RF application technologies	RF component miniaturisation for mm-wave applications Package integration of additional functionality such as antennas, passive devices and power sources Challenges of optical packaging for recent optical sensors (Visible, NIR, IR)	RF miniaturisation for THz applications Packaging of wide bandgap materials (GaN, SiC, including specific challenges for higher power and voltage levels, etc) Optical packaging for PIC, lidars and for hyperspectral imaging applications	New cryogenic compatible packaging platforms for QIP Packaging solutions for extremely high power dissipation and new materials
	Topic 3.3: 3D integration technologies	Chip I/O-pad level 3D-SiC Chip-package-board co-design	3D SoC System technology co-optimisation 3D stacking for quasi-monolithic integration (such as "sequential layering")	Ultimate transistor-level 3D ICs ("monolithic" integration)

Major challenge 4: Advanced wafer fab equipment solutions	**Topic 4.1:** Wafer fabrication equipment	Manufacturing equipment for 2 nm node logic and memory Manufacturing equipment for 3D heterogeneous integration interconnect and interposer concepts down to 1 µm pitch	Manufacturing equipment for 1 nm node logic and memory Equipment to enable novel switches, transistors and alternatives based on, for example, 2D materials, topologic insulator and spin-wave devices Manufacturing equipment for 3D monolithic interconnect and interposer concepts below 1 µm pitch.	Manufacturing equipment for sub-1 nm node logic and memory Manufacturing equipment for sub-1 nm node transistor, 3D monolithic and optical interconnect concepts
	Topic 4.2: Wafer fabrication equipment for differentiated and/or mature technologies	Equipment for manufacturing of components with advanced nanomaterials Production tools for III-V, GaN, SiC or other material substrates of various thickness Production and metrology tools for wafer processing and alignment on opaque wafer substrates	Equipment for materials and processes for new eNVM types such as (high-density) ReRAM Production tools for 300mm for III-V, GaN, SiC or other material substrates of various thickness	
	Topic 4.3: Automated manufacturing technologies	Semantic integration of data for the modelling of finer process interactions over wider fabrication domains AI/ML enhanced systems Smarter detection of potential deviations and faster reaction through AI augmented insight and diagnostics Integrated Supply Chain to optimize production flows and resources utilization.	Autonomous solutions for factory operation (smart control rooms) with >90% of the deviations treated autonomously by the system. Automatic identification and allocation of tasks to responsible "agents" (whether human, robotic or software).	Quantum computing, semantic modelling and ubiquitous connectivity enable next level digital twins to operate 3D SoC ecosystems.
Major challenge 5: Advanced packaging, assembly and test equipment solutions	**Topic 5.1:** Packaging and Assembly equipment	Packaging & assembly equipment for chip-to-wafer stacking, fan- out WLP, multi-die packaging, "2.5D" interposers and TSVs 300 mm photonic SOI 200 mm POI	Packaging & assembly equipment to enable next-generation autonomous sensors, power electronics and RF/optical communication packaged ICs	Assembly equipment for PIC's, waveguides, photonic interposers
	Topic 5.2: Test and inspection equipment	Test & inspection equipment for chip-to-wafer stacking, fan- out WLP, multi-die packaging, "2.5D" interposers and TSVs.	Test & inspection equipment to enable next-generation autonomous sensors, power electronics and RF/optical communication packaged ICs Test and inspection equipment for advanced die to wafer stacking integration, such as nano-TSV, Hybrid Bonding schemes	Test and inspection equipment for photonics testing, active photonics interposer testing, waveguide quality control.
Major challenge 6: Sustainability of semiconductor manufacturing	**Topic 6.1:** Semiconductor manufacturing line	Reduction of CO_2 and GHG emission and of the electrical consumption of semiconductor production lines Use of recycled and reclaimed water	Reduction of CO_2 and GHG emission from semiconductor production lines through better abatement systems and reduction of gas usage and leakage Use of 100% renewable energy sources	No CO_2 and GHG emission by replacing current GHG-related gases (NF_3, PFC…) by alternative chemistries Use of recycled metals to prevent the scarcity of some mineral ores Use of 100% recycled and reclaimed water
	Topic 6.2: Equipment and material suppliers	Improve energy efficiency of semiconductor equipment	Implement re-use of parts/modules for equipment maintenance and servicing Reduce the use of hazardous materials	

1.2

Foundational Technology Layers

COMPONENTS, MODULES AND SYSTEMS INTEGRATION

1.2 Components, Modules and Systems Integration

1.2.1 Scope

Development and production of smart electronic components and systems (ECS) requires physical and functional integration (PFI) of several functionalities into a new physical entity at component, module and system levels (CMS). Therefore, PFI is one of the essential capabilities required to maintain and improve the competitiveness of European industry in the application domains of smart systems. Although in practice PFI is often application-specific, the materials, technologies, manufacturing and development processes that form these domains are generic and should be standardized, interoperable and reconfigurable where possible. Heterogeneous integration of devices and components fabricated with separate and different fabrication processes is key to PFI. This chapter deals with approaches beyond the semiconductor technologies, material families and on-chip integrated systems (SoC) covered in Chapter 1.1.

In the development of ever better smart systems and innovative products, heterogeneous integration becomes more and more important at every level of integration, from semiconductor SoC to System-in-Package (SiP) and ultimately to larger modules and systems. The importance of SiP technologies in integration terms is emphasized by the treatment of SiP in both this chapter and in Chapter 1.1, from their specific points of view. Particularly, alternative technologies (such as additive manufacturing), complementary materials (both at the functional and structural/substrate level) and heterogeneous approaches to assembly, integration and advanced packaging are considered in this chapter. The term heterogeneous integration is used in its widest meaning: a component should be taken to mean any unit, whether individual chiplet/die, MEMS device, passive or active component or assembled package, that is integrated into higher order single components, modules or systems. Developments of heterogeneous integration technologies and platforms include also flexible electronics and photonics solutions. On the other hand, advanced packaging represents the suite of novel technologies, processes and competences that – in a cost-efficient, environmental sound way – allows for the physical, electrical and functional integration of any set of technological diverse components required to build an advanced system in a way that can safely interact with its application environment.

Smart CMS are the key enabling link between basic technologies, e.g. semiconductor or interconnection technology, and key applications as described in the **Application** Chapters. They open the way for widespread use in all application domains by integrating functionalities such as intelligence, sensing, communication and control, even in the smallest devices, through simultaneous development and co-design with **Embedded Software** and **System of Systems** (SoS) technologies and with support from cross-sectional technologies: **Artificial Intelligence** (AI), **Connectivity**, **Architecture and Design**, and **Quality, Reliability, Safety and Cybersecurity**.

The methods, processes and schemes required for the design, production, assembly and testing of the various components, modules and systems and their integration need to be devised with appropriate quality, reliability, repeatability as well as scalability and sustainability (circular economy, CO_2 footprint, life cycle considerations, efficient use of resources).

Considering the new requirements imposed by modern and future smart systems, mastering the integration technologies at CMS levels is a significant capability of European industries. Such a strength needs to be sustained and reinforced to ensure Europe's leading position in smart systems engineering, as well as to bring innovations into real-life reliable and sustainable products, services and markets.

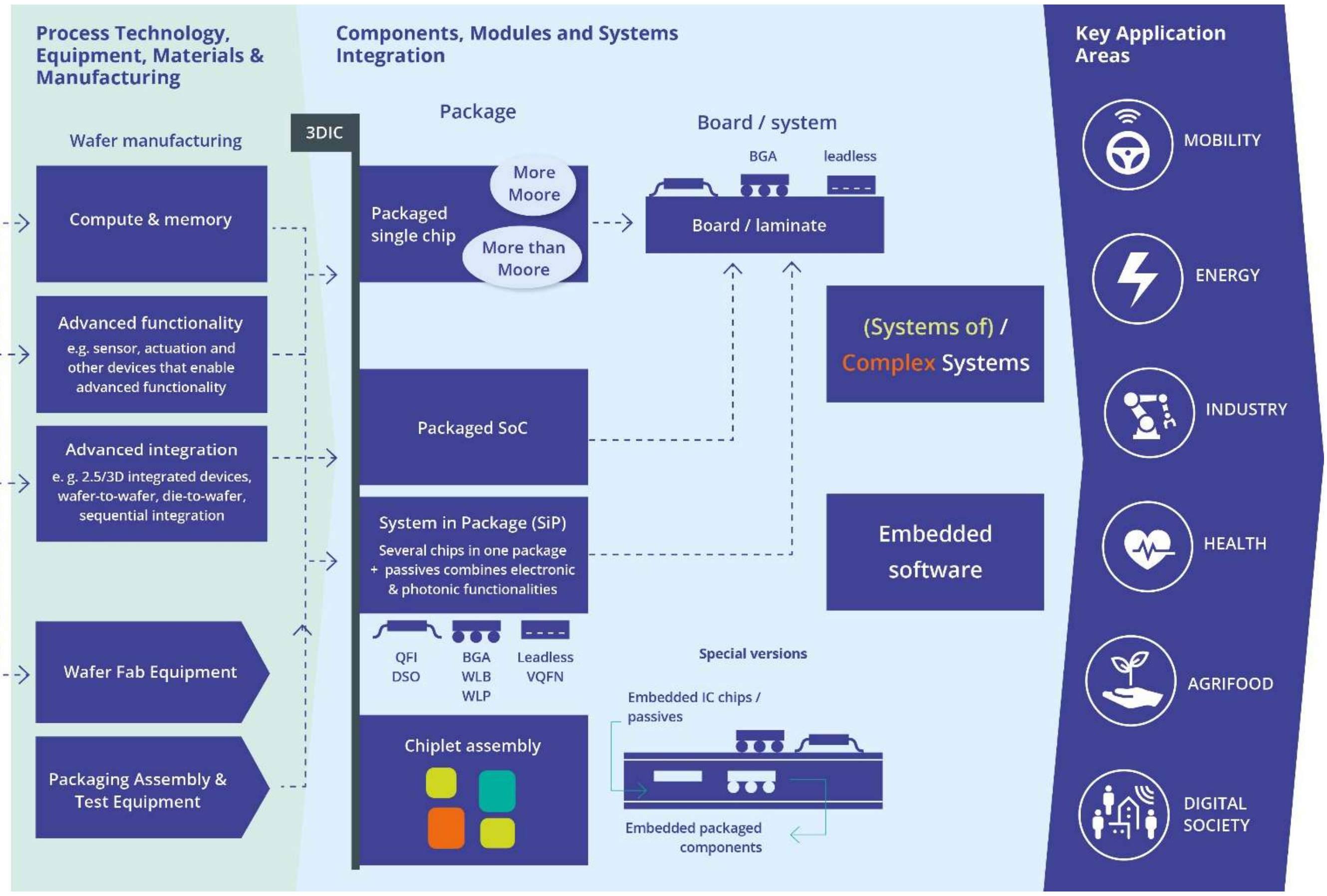


Figure 1.2.1 The Components, Modules and Systems Integration Chapter focuses on physical-functional integration of devices and components into subsystems and systems, using sustainable and efficient materials and integration processes.

1.2.2 Application breakthroughs

Technology advances at CMS level will have a key impact on applications. Future smart CMS will show a strong increase in functional and structural complexity and higher integration levels. They will show more diverse features and material integration in even smaller form factors.

With the ongoing, consistent societal digitization trend in all areas (edge-IOT applications, mobility, telecommunications, medical…), cyber-resilient, long-term available and digital storage of data, OS, firmware and boot media, long operational lifetime and adaptability to variable conditions for ordinary or critical application and infrastructure are coming to the fore.

Applications are the drivers for such approaches:

- Communication landscape with 5G, 6G and increasing data rates, including non-terrestrial networks (NTN), time-sensitive networks (TSN) as well as navigation and localization, including optical integration, components and systems for fiber networks.
- Autonomous systems in mobility, transport, logistics, manufacturing or control of buildings and micro-grids, etc., ensuring faster time response and decreasing the impact of human error.
- Monitoring at remote or difficult to access locations, e.g. structural health monitoring of infrastructure (bridges, tunnels, civil structures) or harsh environments supported by unattended systems, able to operate reliably with power autonomy and decision-making enabled at the edge.
- Healthcare landscape with applications moving towards prevention, diagnostics, therapy and rehabilitation, and life science and pharma domains is moving towards personalized medicine and

regenerative medicine smart multifunctional systems, deployable in the body, on the body or around the body.

- The transition in mobility towards zero-emission power trains, with their significant cost and energy efficiency challenges, including energy systems for high-power charging and/or highly variable and changing conditions.
- Industry 4.0 manufacturing landscape to enable agility and autonomy, as well as energy and resource efficiency, including manufacturing down to lot-size-1.
- Progression to scalable, fault tolerant and ultimately self-monitoring and repairable/re-configurable networks particularly for long life span applications such as environmental and structural health monitoring.
- Sensing of environmental parameters in smart agriculture, livestock or aquaculture, in manufacturing and working places, at home and in urban areas, e.g. for higher yield, energy efficiency and well-being.
- Imaging and audio applications including VR/MR/AR applications for security, healthcare, digital industry, (precision) agriculture, food industry, digital society (television, social media) and perception, requiring deployability, portability and functionality of electronics, photonics and information systems.
- Enabling repair as business, including repair index and set-up of repair processes. This also includes self-monitoring for condition evaluation.

1.2.3 Major Challenges

The following Major Challenges are identified:

- **Major Challenge 1: Functionality.** Developing new features for power, sensors, actuators and smart systems via new materials and methods enabling miniaturization and optimized performance.
- **Major Challenge 2: Advanced integration solutions.** Leading edge and advanced packaging technologies, design, materials, testing, including chiplets and integration on system level.
- **Major Challenge 3: Heterogeneous integration.** System-technology co-optimization for integration technologies, processes and manufacturing enabling cost-competitive smart systems supporting strategical autonomy.
- **Major Challenge 4**: **Sustainability.** Sustainable components, modules and systems and their integration processes to minimize their environmental impact over the entire lifecycle.

1.2.5.1 Major Challenge 1: Functionality

1.2.3.1.1 State of the art

Physical and functional integration (PFI) considers the development of new elements and methods enabling more functionalities to be integrated physically on CMS, in the most effective form factor. This requires interdisciplinary technology innovations as smart CMS may utilize a combination of features based on nano-electronics, micro-electro-mechanic, thermoelectric, magnetic, photonic, micro-fluidic, optical, acoustic, radiation, radio frequency, biological, chemical and quantum principles. Furthermore, many types of devices are to be integrated together, such as sensors, actuators, energy generators, energy storage devices, data processing devices, transceivers and antennae. Different technological approaches such as mainstream silicon technologies, MEMS/NEMS, MOEMS and LAE can be combined for the synergistic assembly of electronic and photonic devices.

PFI goes beyond the compact monolithic SoC approaches supported by the semiconductor technologies covered in the **Process Technology, Equipment, Materials and Manufacturing** Chapter. It addresses System in Package (SiP) realizations involving different integration methods of components with a higher degree of technological heterogeneity into different platforms, including Integrated Circuit carriers, PCB boards, additive manufacturing on flexible or hybrid substrates, and co-packaging of optics and electronics.

PFI requires not only the integration of physical components together, but also the co-design and integration of software, especially embedded software and embedded AI accelerators, to create reliable and sustainable functional systems. This also extends to the development of modular architectures that enable such systems to be configured dynamically and optimally for a given application. However, this chapter concentrates on the physical integration of hardware, including computational devices into systems, while leaving the software edge to the **Embedded Software** Chapter.

1.2.3.1.2 Vision and expected outcome

Given the broad range of physical scenarios they face, smart CMS need to interact with many environments, ranging from lab to industrial and harsh environments to in vivo.

There is a multitude of operational issues affecting CMS regarding energy, performance and size. With respect to energy, for portable IoT devices, there is a need for low-power operation and provision of energy autonomy (self-powered devices or devices providing autonomy adequate to meet the application need (lifetime, service intervals, etc.), depending on the application) on the one hand and dealing with high-power density and thermal stress on the other. Regarding size, the optimal “minimum” size must be achieved. Heterogeneous integration and advanced packaging and interconnection technologies need to be utilized to achieve the best performance in the smallest system-level form factor with the lowest power consumption, or at least take energy availability into account. In high-end consumer electronics, high-performance and ultra-dense compact interfaces at all integration levels are needed, from chiplets and chip assembly and packaging to component and module level connectors and connections with an increasingly widespread photonics communication approach.

MEMS/NEMS/MOEMS development focuses on sensors and actuators that benefit from the free surfaces and volumes that MEMS/NEMS/MOEMS processing is able to produce in a semiconductor substrate. The former relies on new generations of inertial measurement units (accelerometers and gyroscopes) with increased performance, with or without AI support, magnetometers, pressure sensors, microphones, as well as particle sensors; the latter rely on piezoelectrically, electrostatically or electromagnetically driven micro-mirrors, print heads, oscillators (membranes and cantilevers), tunable lenses, loudspeakers and piezoelectric micromachined ultrasound transducers (pMUTs). New piezoelectric materials, such as scandium aluminum nitride (ScAlN) enable new applications and improved performance for MEMS devices, e.g. in acoustic RF filters, in optical systems such as lidars and in ultrasonic sensing. MOEMS sensors such as microbolometer thermal imagers which are MEMS above IC components also rely on the thermal resistance change of thin pixel membranes induced by the absorption of infrared radiation. For hydrogen detection applications, existing MEMS technologies for pressure sensing are used and optimized. Additionally, new MEMS concepts for hydrogen detection are assessed to enable the future hydrogen economy.

The technical challenges for new and future integrated photonic CMS initially lie in a suitable co-design strategy (including associated methods/tools) of the various technologies to be combined as well as in their common processability and hardware integration (compatibility). The high level of integration also requires a special focus on thermal management, optimized thermal design, and suitable cooling concepts. This aspect, as well as other requirements, will determine the design and technology of future packages of photonic components. A particular focus is on those cases in which photonics and electronics have to work together at high bandwidths, such as very high-capacity transceivers, interfaces to electronic switching circuits in communications and data systems, high pixel-count active sensors, 3D imaging and displays. Another technical challenge is the lack of standard solutions for hybrid integration. Additionally, we face a lack of standards for development (e.g. qualification testing) or even legal framework conditions (e.g. for the regular use of smart glasses and similar applications).

Heterogeneous integration technologies are strongly driven by consumer applications, such as the various types of portable and handheld devices. The manufacturers and associated supply chain for these high-volume applications are primarily based in Asia, and so for PFI Europe needs to reinforce its supply chain of integration and packaging solutions. The convergence between sensing and imaging domains for consumer applications, for example face recognition and augmented reality (AR) based on consumer lidar solutions, requires co-integration of (high-speed) electronics and Integrated Photonics into compact systems. Further, portable consumer electronics utilize flexible structures and technologies, which aim at even thinner and more flexible electronic components and systems e.g. for displays, wearables and novel human-machine interfaces (HMI). Structural and 3D electronics enable incorporating electronics in 3D surfaces and mechanical components by means of moulding, additive manufacturing or laser direct structuring (LDS). Novel flexible and stretchable substrates (such as thermoplastic polyurethane (TPU) and polydimethylsiloxane (PDMS)), as well as new materials for active components, including conductive and dielectric inks, with organic materials, metal oxides, nanomaterials and 2D materials are required for realizing new applications from touch panels to RF antennas, control electronics, embedded lighting and sensors/actuators.

One of the key application drivers for the PFI of smart CMS is the IoT and its sensor nodes, which require a wide range of sensor and actuator functionalities, combined with data processing and wireless communication, and with power autonomy provided by energy storage and harvesting devices. Without an adequate alternative to primary batteries, IoT will not meet in full its deployment expectations. This also requires the development of low-power solutions for sensors and actuators, as well as radio communication components and processing. Thermal management challenges introduced by increased functionality in a minimum form factor need to be solved. New and improved energy storage, especially low-leakage rechargeable storage devices, in many cases capable of operating in harsh environments (temperature, ingress, in-vivo) needs to be developed as well as universally deployable and scalable harvesting solutions to improve the case-specific devices used today. In addition to this, it is important to improve in low-power techniques at the system level with co-design of hardware and software and overall system architecture, e.g. in wireless sensor networks, to ensure reliable, sustainable and energy efficient data collection and processing systems attending to the energy available and the possible processing capabilities split at the edge, gateway and cloud. Reliable and fault-tolerant wireless networks are required for applications where long-term continued sensing is critical (e.g. structural health monitoring of civil infrastructure).

Another domain strongly reliant on PFI are Electronic Control Units (ECUs). Due to their complexity and high degree of integration, these systems benefit from advances in generative design, miniaturization, scalability, increased processing power, cybersecurity, AI and machine learning integration, prognostics and health management as well as sensor fusion.

Components to provide power efficient computational resources, i.e. low-power microprocessors and devices with novel computational architectures such as neuromorphic devices, are needed, as are low-power computational methods, including distributed and low-power AI solutions in hardware, software, and in-sensor and data fusion data processing. In addition, reliable, energy-efficient, scalable, low-loss interconnection and packaging solutions are a necessity.

Smart CMS leverage a multitude of materials, such as silicon and other-than-silicon semiconductors, precious and rare earth metals, ceramics, polymers, glass, inks and functional materials for sensing, actuation and energy harvesting, as well as hybrid combinations of substrates and materials (e.g. Si, ceramic, polymer, glass, metallic glass), in packages and in systems, extending the coverage of the usual materials in semiconductor-based technologies. Many of the new features required by future smart systems can only be achieved by introducing novel materials into the devices and systems, from back-end of the line processing of the microchips, or by post-processing on CMOS, to novel IC carrier technologies, including fan-out wafer level packaging and other SiP technologies as well as PCBs or in printing, additive manufacturing or other means. The development of new materials and the compatibility of those (with regard to e.g. process compatibility, environmental compatibility) is critical to the future development of PFI.

Quantum technology provides a new modality for More-than-Moore technologies not industrially considered earlier due to technology limitations in field-deployability, extreme cooling requirements etc. Recent years have seen rapid technological developments in the technological readiness of many quantum technologies: Quantum sensors, especially those based on gas cells, are being industrialized and miniaturized, promising better stability and accuracy of measurements, e.g. in magnetometry and inertial measurements. Quantum computers are developing into larger and more powerful machines year after year. Quantum communications, or quantum key distribution (QKD), enable secure encryption based on quantum states of photons. The rise in the applicability of quantum technology is based on successes in materials, fabrication processes and quantum science itself, but also in major part in the successful development of enabling technologies, such as control and low-noise readout electronics, packaging and heterogeneous integration and cryogenic cooling solutions, many of which are based on the successes and processes of electronics and semiconductor industry. Hence, the further development of these enabling technologies should be considered in detail.

The following key focus areas address the multimodality of ECS, which goes beyond semiconductor technology, requiring advanced packaging and heterogeneous integration of diverse materials, components and platforms. Full coverage of the physical-functional integration requires considering both the physical integration technology platforms and the functionalities of the integrated systems.

1.2.3.1.3 Key focus areas

- Sensing, imaging and actuation:
    - Sensors and actuators leveraging the integration of MEMS/NEMS, MOEMS and micro-optics elements.
    - Sensors and actuators for biological, medical and diagnostic applications, and for sensing of human vital signs and biomarkers, as well as for selective detection of gas and volatiles, allergens, residues/pollution in food/water, atmospheric particles, hazardous substances and radiation.
    - Sensors and systems enabling integration behind OLED and sensing through the OLED display.
    - Smart surfaces for HMI including integrated sensors for micro-LED and mini-LED displays.
    - Sensors and systems for sensing and imaging in the short wavelength infrared range, based on Ge on Si Integration or Pb-free quantum dots, enabling a broad range of new applications in the areas of lighting, biotechnology and life sciences, photovoltaics and information processing.
    - Advanced global shutter and rolling shutter CMOS Image Sensors (CIS) based on novel pixel technology for VIS-NIR imaging applications, with the goal to improve the performance, sensitivity, and efficiency of VIS-NIR CIS pixels, enabling high-quality imaging across a broad spectral range, as well as infrared imaging.
    - Imaging systems: lidars and radars, including multi-modal and hyperspectral, i.e. spectrally resolved, sensors.
    - Sensors and systems utilizing quantum principles, e.g. single photon sensors, including required cryogenic and cooling components and systems.

- Devices with new features and improved performance for sensing and actuation using novel materials (metal nanowires, carbon nanotubes (CNTs), graphene and other 2D materials, cellulose nanofibers, nitrogen vacancies in diamond, metamaterials, metallic glass etc.), in combination with, or integrated on, CMOS.
- Sensor fusion and virtual sensors including appropriate data and communication infrastructure, e.g. for condition monitoring, prognostics and health management for ECUs.
- Ultra-low power event-based sensors, e.g. inertial motion & image detection for asset tracking, incident/anomaly detection, etc.
- Dedicated LED and laser light sources from UV to visible to IR for sensing applications by avoidance/replacement of hazardous material (e.g. Cd/Pb-free QDs for color conversion, UV-C LEDs instead of Hg-lamps).
- Materials that ensure the hermetic sealing of sensors, actuators and systems and at the same time contribute to the miniaturization of components.

- MEMS technology
    - CMOS or GaN-compatible thin film piezoelectric materials, such as ScAlN for piezo-actuated MEMS sensors and actuators.
    - Acoustic piezo-MEMS devices, pMUTs, and acoustic RF filters for high frequencies above 6 GHz.
    - Audio MEMS systems such as micro speakers with more advanced integrated functionality, such as amplifier and integrated noise canceling (ANC).

- Integrated Photonics
    - New materials for active photonic devices, such as 2D materials, Lithium Niobate, Indium Phosphide; for improved performance, such as higher bandwidth in modulators, and detectors.
    - Light sources (e.g. µLEDs, lasers and laser modules) with higher power and better performance and with tunable wavelength, using external cavity on photonic integrated circuits (PIC). Due to the high level of integration, electronic and photonic components must not be treated independently. Development of such components goes beyond the scope of the ECS-SRIA and is handled by Core Technologies Working Group of Photonics21.
    - Microbolometer thermal imagers with integrated optics and edge AI recognition for safety and mobility applications such as Automated Emergency Braking (AEB) and sensing for autonomous vehicles. Pixelated light sources for illumination, visualization and communication purposes including vertical hetero-integration of III-V materials on Silicon.
    - New waveguide materials and components to expand the wavelength range from UV up to mid IR optical elements for beam shaping and manipulation (like ultra-thin curved waveguides, meta-lenses, tunable lenses and filters, next generation holograms, ultra-wide-angle holograms).
    - Display technologies (like OLED-on-silicon, micro-LEDs, MEMS-mirrors, Phase parrays) and sensors (e.g. for eye tracking).
    - New devices for Quantum PICs.

- Flexible electronics
    - Sensing devices and power sources compliant with hybrid integration in wearables, considering flexibility, durability (e.g. washability) and biocompatibility.
    - Flexible and stretchable sensors and modules, e.g. OLED displays, OPVs, touch surfaces and other sensors/actuators, conformal antennas.

    - Functional materials for flexible and stretchable devices; organic and inorganic semiconductor materials and inks, perovskites for OPV.
    - Barrier materials, dielectrics, and transparent conductor materials and inks for flexible electronics and additive manufacturing.

- Communications
    - Module-level high-speed wireless communication features, including current and new frequency bands.
    - High-speed photonics communications modules beyond 1 Tb/s.
    - New front-end components, filters and functionalities, e.g. active antennas for 5G and 6G communications and non-terrestrial network solutions.
    - Low latency and low power communications in-package/module as well as at system level for the edge and IoT devices.
    - Energy and application context aware comms that can dynamically adapt the architecture (edge, gateway, cloud) to meet sensing, processing and actuation needs based on limited energy available, especially at the edge. Continuous delivery of new features and fixes through Over-The-Air (OTA) updates to ensure the security of the device over time and reduce the digital waste by increasing the life span of a device.
    - Strategies and components for Electromagnetic interference (EMI) mitigation and reliable operation in harsh environmental conditions.

- Energy and thermal management
    - Low-power/low-loss modules for low-power sensing, actuation, processing and communication.
    - Multi-source power architectures with digital interfaces driven by dynamic and context aware algorithms that can adapt based on energy available versus needed for sensing, actuation, processing and communication.
    - Energy-autonomous multi-sensor modules and systems including energy harvesting, sensing, actuation, processing and communication.
    - Power management components and modules compatible with harsh environments (high temperatures, vibrations, bio-compatibility, ingress, electromagnetic interference (EMI) conditions for industrial, automotive, med-tech and space technology).
    - Devices using non-toxic materials for efficient energy sources, storage and harvesting devices (thermoelectric, piezoelectric, tribo-electricity, etc.), and higher performing electrodes and electrolytes for improved capacity and low leakage of energy storage devices or new lightweight energy harvesters for mobility and transportation applications.
    - Solutions for thermal management for integrated photonics and RF systems at different integration levels including advanced and active cooling systems.
    - Thermal management and smart cooling systems for industrial applications and harsh environments.
    - Efficient smart compact cooling solutions and approaches for quantum devices and cryogenic multiplexing with semiconductors or superconducting devices.

- Information processing
    - Component and system-level features for self-diagnosis and module-level signal processing and control features for self-diagnosis, self-monitoring, and self-learning and self-repair.
    - Sensor level hardware and software solutions for security and privacy and data reliability.
    - Machine learning and artificial intelligence and data analysis at the sensor, module and systems level, i.e. on the edge data analysis embedded at different levels for smarter devices, including AI at sensor level.
    - Integrated and scalable solutions, both Software and Hardware integration with increased processing power for more sophisticated features, especially for edge AI and ECUs.

- Use of quantum computing and Integrating quantum computing for data-analysis.

- Integration of electronic components in electronic control units:
  Another domain strongly reliant on PFI are Electronic Control Units (ECUs). Due to their complexity and high degree of integration, these systems benefit from advances in generative design, miniaturization, scalability, increased processing power, cybersecurity, AI and machine learning integration, prognostics and health management as well as sensor fusion.
  - Sensing, imaging and actuation
    Sensor fusion and virtual sensors including appropriate data and communication infrastructure, e.g. for condition monitoring, prognostics and health management for ECUs.
  - Information processing
    Integrated and scalable solutions, both Software and Hardware integration with increased processing power for more sophisticated features, especially for edge AI and ECUs
  - Advanced packaging and SiP technologies
    Robust heterogeneous 3D integration of sensors, actuators, electronics, processing units, communication, RF front-end components and energy supply into miniaturized systems.

1.2.5.2 Major Challenge 2: Advanced integration solutions

The major challenge Advanced integration solutions aims to address the growing demand for smaller, more efficient, and reliable electronic components. This challenge will focus on developing advanced packaging technologies that enable the integration of multiple functions and components into a single package, while also improving thermal management and reliability. Key areas of research will include advanced materials and manufacturing processes, as well as the development of new design and simulation tools to optimize the performance of integrated electronic packages. By fostering collaboration between industry, academia, and research institutions, this agenda seeks to position Europe as a global leader in advanced electronic packaging solutions.

1.2.3.1.4 State of the art

The state of the art in electronic packaging worldwide is characterized by a strong emphasis on miniaturization, integration, and reliability. Globally, there is a significant focus on developing advanced packaging technologies that enable the integration of multiple functions and components into smaller and more efficient packages. This includes the use of advanced materials, such as organic, silicon or glass substrates and advanced interconnect technologies, as well as the development of innovative manufacturing processes, such as 3D integration and wafer-level packaging. Main research focus worldwide is on chiplets technology. Additionally, there is a growing emphasis on improving thermal management and reliability through the use of advanced simulation and testing methods.

The European industry is actively involved in revitalizing electronic packaging technologies and is a key player in the global landscape, especially in the field of electronic packaging for harsh environments. European companies and research institutions are at the forefront of developing innovative packaging solutions for power electronics, contributing to the overall state of the art in electronic packaging worldwide. Further collaborative efforts and research initiatives are required, to keep the European industry in leadership position.

1.2.3.1.5 Vision and expected outcome

The Advanced integration solutions challenge envisions a future where electronic devices are smaller, more efficient, and reliable. The expected outcome is to establish Europe as a global leader in advanced packaging technologies, with a focus on miniaturization, integration, and reliability. Through initiatives like the Pack4EU project, the vision is to develop innovative packaging solutions that enable the seamless integration of multiple functions and components into smaller and more efficient packages. This will be achieved through 9 recommendations (source Pack4EU):

- Implementation of an **Industrial Transfer Instrument** dedicated to semiconductor packaging bridging the gap between prototyping and production.
- Establishment of a **High-Level Packaging Board** composed of industry leaders to guide strategic directions driven by market needs, ensuring innovation in packaging standards and practices.
- Implementation of a T**echnical Expert Group** for developing and updating a **Roadmap for Advanced Packaging** in Europe.
- Creation of **Open Piloting Facilities** for small and medium volume production as a seed for growing European Advanced Packaging capabilities.
- Development of tools and methodologies for a **Design-to-X** approach & Standardization for an OPEN European Co-Design Ecosystem.
- Consolidation of international relationships through open calls on **Sustainable Packaging Materials and Substrates**, for example in liaison with IAM4EU and ERMA.
- **A strategy for training and skills development** including: Bridging research and education and facilitating international exchange on training through an EU Education Hub and Building a framework for access to EU-funded research pilot lines for students.
- Support for **Small and Medium Enterprises** through dedicated open calls applying, for example, cascade funding schemes or other funding instruments dedicated/adapted to SMEs.
- Creation of a **Pan-European Network** for Advanced Packaging to federate and strengthen the European ecosystem.

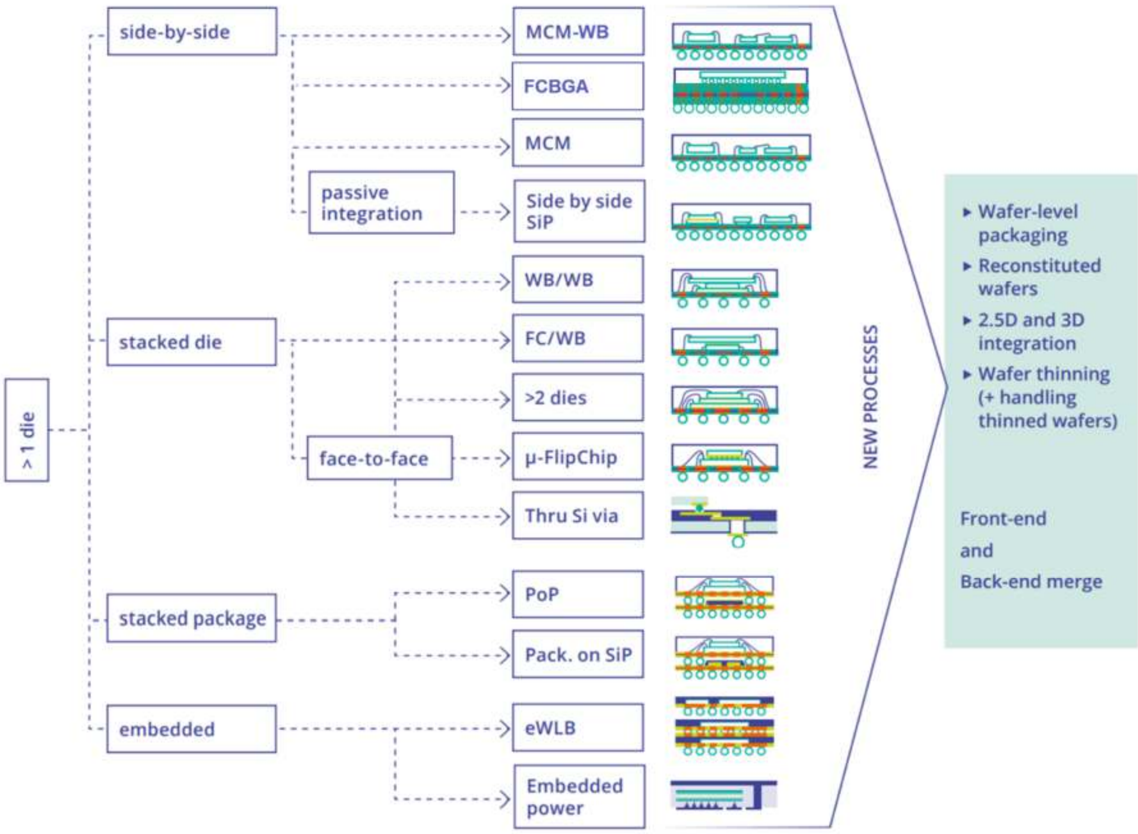


Figure 1.2.2 System in a package (SiP) examples. With merging of front-end and back-end, almost unlimited possibilities enabled (Source: Andreas Grassmann – Infineon – 3D Systems Summit)

1.2.3.1.6 Key focus areas

Research and development priorities are focused on innovative approaches, such as the following.

Advanced packaging and SiP technologies

- Robust heterogeneous 3D integration of light sources, sensors, actuators, electronics, processing units, communication devices, RF front-end components and energy supply into miniaturized systems, often working in harsh environments.
- Embedding of power sources (energy harvesting transducers, batteries, supercaps, etc.) into a package (PwrSiP).
  - Increasing functionality in IC Substrates for high-efficiency power delivery including voltage regulator circuitry and integrated capacitances and other passives. The solutions can entail specific passives dice as part of a SiP. Additionally, fine structuring capability achieving high bandwidth interconnections has to be reached either sequentially (2.1D) or recombining separated manufactured carriers (2.3D and 2.5D).
- IC carriers with integrated voltage regulator and capacitance increasing the power delivery efficiency, as well as finer structuring in IC carriers: below 5/5µm line width and spacing and finer micro-vias (below 15µm diam.).
- Chiplet integration concepts
  - Multi-node chiplets for compute applications involving high speed chip-chip interconnections with high resolution electrical and optical routing in substrate or redistribution.

- Heterogeneous Integration of MEMS, Quantum Computing, Power, etc. into Chiplet-based systems.
- High density (HD) substrates, incl. HD routing components as bridges as enabler for successful chiplet-based systems.
- Low stress assembly using low temperature assembly/interconnections processes, for example nanowires.
- Virtual prototyping methods to ensure low warpage design of the chiplets with stress-minimization.

- Integration with biological and molecular systems, including fluidics and surface coatings and functionalization materials and methods for multi-functionality on the same base structures, e.g. biosensor arrays on Silicon.
- New functional materials for packaging that enable integration of sensing or other functionality or enhanced functionality into the packaging itself, e.g. packaging as a part of the antenna or sensor functionality, e.g. embedded heatsinking/spreading to increase the thermoelectric energy harvested.
- Customized materials and methods for housings and coating features and new substrate materials for specific requirements: high-efficiency and high-power density, high frequencies, disposable, bio-compatible, non-fossil, harsh environments.
- Rapid prototyping and manufacturing technologies (additive manufacturing, 2D and 3D additive technologies, etc.).
- High-performance materials for passives enabling close coupled passives for high-density heterogeneous integration such as magnetic cores, high-k dielectrics.
- Ultra-dense and small interfaces in all integration levels, from chip assembly and packaging to component and module level connectors and connections.
- Integration of different sensors and sensor hubs for sensor fusion (e.g. combination of acceleration sensor, microphone, micro speaker for enhanced noise cancellation).
- Manufacturing and characterization processes for hermetic sealing of components or subsystems with low leakage level.
- New failure analysis tools and methodologies including digitalization methods such as digital twin, AI/ML to improve efficiency and robustness for leading edge and advanced packaging and shorten development time.
- Innovative equipment and methodologies supported by digitalization methods for pre-assembly of thin silicon dies to improve production quality.
- Interconnect technologies that allow 3D stacking as well as horizontal interconnecting of dice and chiplets. This also includes interconnects through optical interfaces, most notably off-chip, but also within a package.
- Packaging for RF application technologies:
  - Solutions for high-frequency miniaturisation, such as for mm-wave applications (> 60 GHz) and for > 100 GHz, including antennae integration THz applications for which no package solutions currently exist.
- Packaging for photonic application technologies:
  - Solutions for advanced optical functionalities in the packages, either for sensors (visible, NIR, infrared), or for PIC.
  - Specific solutions for the signal conditioning of other sensors (mechanical, physical, chemical…).
- Packaging for power discrete packages and power modules:
  - High-Tg and high-thermally-stable polymer materials for packaging and high-performance thermal interface materials that will enable more efficient operation of WBG devices
  - Integration of power devices with logic and sensors in one package.

- Enhanced reliability, robustness, and sustainability technologies:
  - Solutions for high reliability, robustness and high quality taking into account packaging materials and related interfaces. For this, a close consideration of the chip/package interaction, but also of the interaction of chip/package to the board, is required. R&D in this area requires a strong link, especially with materials and their compatibility, and consideration of the heat dissipation challenges. In addition, variations and extremities in operating environmental conditions should be considered to ensure devices work seamlessly and operational life is not impaired. Avoiding (particle) contamination is another, increasingly critical, requirement. In the last decade nearly all assembly and packaging materials have changed; in the next 10 years, it is expected they will change again. Also, a close link with the Architecture and Design section is crucial here.
  - Solutions to test separate components, before and after assembling these in a single package/ subsystem. Concepts like built-in self-test (BIST) and self-repair require some amount of logic integration, and a design providing access for die testing.

System requirements and semiconductor device technology (Major Challenges 1 and 2) will evolve at the same time, creating momentum for further interconnect pitch scaling for 3D integration technology platforms. Hence, the timelines of all four challenges of this section are strongly connected.

1.2.5.3 Major Challenge 3: Heterogeneous integration

1.2.5.3.1 State of the art

Smart CMS require a multitude of processes: silicon and other micro- and nano-processing, additive manufacturing, lamination and other interconnection and assembly technologies, as well as hybrid combinations. To increase the integration density and combine the above-mentioned features, many different integration and packaging technologies are required, such as thin film processes, embedding, classic assembly and joining methods, both for single components as well as modules.

Heterogeneous integration, from components up to the system level, requires-engineering on many technology domains, such as power, signal integrity, EMC, thermal and mechanical. All such domains and their hardware and software interplay (power and communications) must be designed together to ensure a high device- and system-level performance and the necessary integration. The challenge here is to combine all these domains in the design and simulation of integrated systems, often with inadequate information of all the properties of the included materials, components and processes. Where possible historical data and related algorithms should be used to predict and optimize system level performance. This may involve changing sensing intervals and/or how and where data is processed and routed. Furthermore, standardized and interoperable design and simulation methods that enable and support such multi-physics and multimodal design and manufacturing must be addressed, with the possibility to parameterize not only the material parameters, but also the system parameters, e.g. variability in the quality of the contact (e.g. thermal, mechanical) between the transducer and the ambient energy sources in the case of power harvesting. Modelling and design tools for thermal, mechanical and electrical characteristics in small 3D packages, including moulded and additive manufacturing methods are needed, linking to the **Architecture and Design; Methods and Tools** Chapter.

Heterogeneous chiplet integration enables the creation of highly customized, optimized and advanced systems by leveraging specialized chiplets for different functions. It enables faster and cost-effective manufacturing, as smaller chiplets have higher yields and can be produced independently. Furthermore, the ability to integrate chiplets from different manufacturers cultivate a more competitive, flexible and dynamic ecosystem. Designing systems with multiple specialized chiplets requires advanced design tools and methods to ensure the seamless functioning together. Each chiplet may have different power, performance and communication requirements, necessitating careful planning and integration (hardware and software). Developing cutting-edge interconnect solutions such as silicon, glass, and organic substrates-based interposers, through vias, and micro-bumping to enable high-density, high-bandwidth connections between chiplets. Thermal management is another key challenge, as densely packaged chiplets can generate substantial heat. Developing innovative cooling solutions and advanced materials is required to efficiently dissipate the generated heat to maintain performance and reliability.

Flexible electronics is an enabler to reduce the weight, volume and complexity of integrated systems and products, to create novel form factors and 3D design features. Currently, the majority of flexible electronics products are based on polyimide (PI), copper laminate substrates, etching of copper to pattern the circuitry, and conventional SnAgCu (SAC) soldering or anisotropic conductive adhesives (ACA) bonding processes for the assembly of discrete components on the substrate. Development towards smaller feature size in printing technologies increases the requirement of registration, or layer-to-layer alignment accuracy. Development of IC interconnection and bonding technologies, especially to flexible and stretchable substrates, is critical for improved performance, yield and reliability. In addition, pilot lines and fabrication facilities and capacities need to be developed.

In integrated photonics, if it comes to monolithic integration, the optical waveguides and devices are fabricated as an integrated structure onto the surface of a substrate, typically a silicon wafer. As a result of integration, complex photonic integrated circuits (PICs) can process and transmit light in similar ways to how electronic integrated circuits process and transmit electronic signals. New waveguide and active materials are constantly developed, e.g. SiN waveguides on silicon or Ge detectors and 2D materials for active components, such as modulators or detectors. The fabrication of PICs consists of a multi-faceted integration problem, including monolithic integration, heterogeneous integration of active components (e.g. laser sources), and high-speed driving electronics for e.g. high-speed communications above 100 GHz bandwidth, thermal management and other functionalities, such as fluidic functions for bio and medical sensing. Further system development regarding integrated photonics includes an often 3D assembly of electronics and photonics with passive optical MEMS as well as optics components like lenses, mirrors and beam splitters.

Integration and packaging technologies for quantum systems are key enablers to make quantum sensors and other systems industrially applicable. The integration and packaging in quantum technology poses several non-typical requirements for preserving the quantum coherence. These requirements include use of non-typical materials, such as extremely-low loss dielectrics and superconductors. Further, the packaging must support extreme cooling or even cryogenic operation, either by integrating a cooler technology inside, or by being inserted in an external cryo-cooler, which also entails vacuum operation. Thermal conductivity and handling of thermal expansion of different materials at every level, from wafer and chip level integration methods, to packaging and connection to room temperature are critical. Especially quantum computing requires multi-channel high-fidelity control and read-out solutions, with accurate synchronization and timing, from low-frequency to GHz range to the optical range of frequencies, depending on quantum system modality. Cryogenic electronics, cryo-CMOS or similar, are developed for solving these requirements, at the same time increasing the integration level of the quantum system, by introducing the control and readout elements closer to the qubits or other quantum devices.

The critical requirements to enable new advanced applications are to ensure sustainable and cost-efficient manufacturing while providing optimal performance and reliability. Further important developments include integration of different silicon IC components into miniaturized multifunctional modules following different SiP approaches, combining technologies such as flip chip, bonding, lamination and substrate materials such as silicon, glass, ceramics and polymers. Multifunctional integration also requires the development of multi-domain integration – e.g. the integration of photonic and RF functionalities into smaller form factors and together with sensors and CPUs.

For many portable devices e.g. Wireless Sensors Network (WSN) edge nodes, the power source in itself becomes a more complex challenge to integrate materials and devices, moving from traditional batteries and capacitors. This requires a complex combination of energy harvesting transducers, primary and rechargeable storage devices that need to interact with PMICs, MCUs, sensors and transceivers. Collectively, they need to make decisions at a node and network (gateway, cloud) level on when to use versus store energy and where to take it from based on energy available. The physical mounting of transducers and electrically controlling their characteristics (e.g. impedance matching) will also be critical to maximize their performance.

1.2.5.3.2 Vision and expected outcome

The challenge of integration processes, technologies and the manufacturing of smart CMS is mainly about dealing with the complexity of heterogeneous integration and scalable manufacturing technologies with different economy of scale approaches. These include "intensive" Si-like technologies, or "extensive" printing-like technologies, which under different assumptions and processing paradigms can offer cost affordability and production scalability. Apart from high-volume applications such as medical patches and RF front-end modules for 5G/6G small cells, many industrial applications may require the availability of CMS in relatively small quantities over decades, which adds a new challenge to the scalability of manufacturing and implementation of the latest technologies.

The complexity and diversity of heterogeneous CMS substantially exceeds that of mere microelectronic components due to their multi-physics and multiple domain nature. In addition, the packages will include integrated functionalities, rather than being "passive" boards and frames. Integration and packaging methods should not compromise, but guarantee and even increase the performance of the interlaying components. Especially low-loss integration methods to enable integration of large RF systems or integrated photonics. These technologies would enable e.g. RF front-ends or active antennas for millimeter wave frequencies enabling novel beyond-5G telecom solutions, or in integrated photonic communication and sensing systems. Merging different PIC technologies to compensate for missing functionality in individual cases is based on a variety of technologies, such as die flip-chipping, wafer bonding, micro-transfer printing, edge coupling, and others. Amongst current examples, one that is particularly pronounced is the integration of III-V light sources with silicon photonics, which is missing from silicon technology on its own.

In this multifunctional and multimodal integration at CMS level, the development of manufacturing methods that meet the accuracy and repeatability criteria of high-quality and high-reliability products for a broad range of applications and constraints (physical, mechanical, thermal, environmental) is challenging and needs development. This method development shall be accompanied by process modelling leading to a digital twin for manufacturing allowing documentation, simulation and improvement of manufacturing challenges in a digital environment (see also Chapter 2.4).

Additive manufacturing can provide structural and functional solutions for smart CMS integration that are not feasible with traditional methods. These methods will enable zero-defect manufacturing starting at lot one. Although additive manufacturing also improves manufacturing flexibility, solutions for the cost-efficient scaling of these fabrication methods must be addressed. 3D CMS integration methods will need to be developed to provide greater functionality and miniaturization in a cost-effective, sustainable and scalable way.

With respect to this multi-modality of heterogeneous integration methods, the key focus areas are divided based on the main technologies: advanced packaging and SiP technologies, integrated photonics, flexible electronics, quantum systems and manufacturing methodology, as follows.

1.2.5.3.3 Key focus areas

- Advanced packaging and SiP technologies
    - Robust heterogeneous 3D integration of light sources, sensors, actuators, electronics, processing units, communication, RF front-end components and energy supply into miniaturized systems, often working in harsh environments.
    - Embedding of power sources (energy harvesting transducers, batteries, supercaps, etc.) into a package (PwrSiP) and on a chip (PwrSoC).
    - IC carriers with integrated voltage regulator and capacitance increasing the power delivery efficiency, as well as finer structuring in IC carriers: below 5/5µm line width and spacing and finer micro-vias (below 15µm diam.).
    - Multi-node chiplets for compute applications involving high-speed chip-chip interconnections with high-resolution electrical and optical routing in substrate or redistribution.
    - Embedding of power sources, ASIC drivers and audio MEMS sensor and actuators such as microphones and micro speakers into a heterogeneous SiP.
    - Integration with biological and molecular systems, including fluidics and surface coatings and functionalization materials and methods for multi-functionality on the same base structures, e.g. biosensor arrays on Silicon.
    - New functional materials for packaging that enable integration of sensing or other functionality or enhanced functionality into the packaging itself, e.g. packaging as a part of the antenna or sensor functionality, e.g. embedded heatsinking/spreading to increase the thermoelectric energy harvested.
    - New materials and methods for housings and coating features and new substrate materials for specific requirements: high power, high frequencies, disposable, bio-compatible, non-fossil, harsh environments.
    - Rapid prototyping and manufacturing technologies (additive manufacturing, 2D and 3D additive technologies, etc.).
    - High-performance materials for passives enabling close coupled passives for high-density heterogeneous integration such as magnetic cores, high-k dielectrics.
    - Ultra-dense and small interfaces in all integration levels, from chip assembly and packaging to component and module level connectors and connections.
    - Integration of different sensors and sensor hubs for sensor fusion (e.g. combination of acceleration sensor, microphone, microspeaker for enhanced noise cancellation).
    - Manufacturing and characterization processes for hermetic sealing of components or sub-systems with low leakage level.

- Integrated Photonics and co-integration with electronics
    - Photonic-electronic system integration based on integrated photonics, including high-speed RF electronics, MEMS/NEMS/MOEMS sensors, etc.
    - Multi-domain electro-photonic integration and electro-optic co-packaging.
    - Wafer-level integration of photonic and electronic components for smart emitters and detectors.
    - Enabling electronic-photonic systems by heterogeneous integration of active components on PICs (III-V semiconductors, ferroelectrics, ultra-low-loss waveguide materials). Heterogeneous integration processes and equipment for integrated photonics, including high-precision component placement and bonding, integration of active components on PICs (III-V semiconductors, ferroelectrics, ultra-low-loss waveguide materials) as well as low-loss fiber coupling to PICs.
    - Quantum PICs: Integration of single photon detectors and sources and quantum photonic systems in PICs.

- Flexible electronics
    - Integration towards low vertical form factor (<100 µm) and the miniaturization of external matching networks through integration.
    - Submicron LAE fabrication processes and equipment (printing technologies in general, nanoimprinting, reverse offset printing, etc.) and automated manufacturing equipment for flexible electronics, including testing tools for electrical and non-electrical properties.
    - Interconnections processes and tools for flexible and stretchable devices and structural electronics (in glass, plastics, laminates, etc.).
    - New/alternative non-fossil, organic, biocompatible and compostable substrate materials, e.g. for implants, ingestibles, wearables, biosensors.
    - Adhesives, bonding materials and methods for integrating chips, light sources, drivers and sensors on flexible substrates.
    - Use of flexible Si-substrates for 3D form factors and for flexible electronics.
    - Integration and embedding of diverse materials and components such as antennae, PV panels, energy storage devices, magnetics, interconnect, heatsinks, displays, etc. in flexible or conformal electronics.

- Quantum systems

    - Materials and methods for integration and packaging of semiconductor electronics, Integrated Photonics and superconducting devices at cryogenic temperatures, including 3D technologies.
    - Integration and interfacing and cabling solutions for combining room temperature systems and cryogenic quantum components, sensors and systems.
    - Integration methods that enable scaling quantum systems efficiently, from wafer level 3D integration to module and system level.
    - Cryogenic electronics, cryo-CMOS and similar for increasing the integration level of quantum systems and enabling scaling up of quantum systems.
    - Development and miniaturization of cryogenic cooling systems, e.g. solid-state coolers.
    - Quantum PICs: Integration of single photon detectors and sources and quantum photonic systems in PICs.

- Manufacturing methodology, characterization and testing
    - Automation and customization in CMS integration for large-scale manufacturing, including Industry 4.0 techniques, design for manufacturing based on production data techniques, and lot-size-1 manufacturing, e.g. BIST (built in self-test) capability of components.
    - Manufacturing and testing tools (including tests, inspection) for CMS, enabling zero-defect integration.
    - Process modelling approaches with focus on productivity, yield, trustability and distributed manufacturing.
    - Material properties database for simulation and reliability based on a standardized ontology.
    - Design of new materials from properties requirements by the means of Materials by Design, materials genome and digital design approach.
    - Energy-effective joining methods, e.g. low-temperature soldering; selective heating processes (i.e. inductive or reactive) to limit overall temperature impact to the system while packaging.
    - Self-powered embedded sensors for ongoing performance and condition monitoring and tracking of devices, e.g. for provenance, life cycle assessments, field failure analysis and authentic validation.

1.2.5.4 Major Challenge 4: Sustainability

1.2.5.4.1 State of the art

In 2019, world-wide e-waste exceeded 50 million tons and is forecasted to grow to 70 million tons in 2030[1]. The European Union is one of the most advanced actors in e-waste recirculation processes. Indeed, in this region 42.5% of e-waste is documented to be collected and recirculated as product or component or recycled as material, whereas this is the case for only 9.4% in Americas and 11.7% in Asia[2]. This can be explained by the clearer European and national political support for such initiatives.

However, as increased integration will cause the borders between CMS to become blurred, and more diverse and complex materials are used at each level, the dismantling of systems into their constituent components at the end of their useful life will become increasingly difficult. Many industrial ECS products have lifetimes extending to decades, thus the environmental regulations for recyclability cannot be known in detail by the time of product design. Nevertheless, early consciousness on the issue should preside the start of the product cycle. Based on identified challenges, regulatory measures under the Eco-design Directive are intended to establish design for energy efficiency and durability, repairability, upgradability, maintenance, reuse, repurposing and recycling. The European Commission presented its new **Circular Economy Action Plan** (CEAP) in 2020 to limit waste generation and encourage recycling, product repair and reuse. Part of this initiative is the digital product passport (DPP), which informs end-customers and businesses about products' sustainability. In general, the concept of the circular economy (CE) builds upon well-implemented strategies to prevent waste generation and includes eco-design rules during product design phase and measures for re-using products and components after a use phase.

The European Platform on Life Cycle Assessment (EPLCA) includes information on the Product Environmental Footprint (PEF) and Organization Environmental Footprint (OEF) methods as a common way of measuring environmental performance (EU Commission Recommendation 2021/2279). The PEF and OEF are the EU's recommended LCA-based methods to quantify the environmental impacts of products (goods or services) and organizations.

ECS should produce smart systems not only as an enabler for, but also as an element of the circular economy, considering the sustainability of the ECS value chain and the products themselves. Focus should be on the sustainability of the CMS production, including processes, materials and maintenance during the primary lifetime. The recyclability of the product must already be considered in the design and manufacturing phase to enable repairability, upgradeability, reconfigurability, extension of lifetime, and repurposing/re-use in a second life application, and finally the recovery of components and materials for recycling.

Given the increasing burden of improperly dealt with e-waste and considering that a significant part of CO_2 emissions arises from the fabrication of the ECS themselves, extending product lifetime is important for reducing ECS-related environmental load. This needs to be addressed by designs that enable repair or replacement of faulty components, avoiding the replacement of the full module or system. To fight obsolescence, hardware and software upgrades should be supported, even in field conditions. Reducing CO_2

---

[1] Forti V., Baldé C.P., Kuehr R., Bel G. The Global E-waste Monitor 2020: Quantities, flows and the circular economy potential. United Nations University (UNU)/United Nations Institute for Training and Research (UNITAR) – co-hosted SCYCLE Programme, International Telecommunication Union (ITU) & International Solid Waste Association (ISWA), Bonn/Geneva/Rotterdam.

[2] NU/UNITAR SCYCLE – Nienke Haccoû

emissions during the lifetime of the system requires minimizing the power consumption at CMS levels while in operation by using low-power hardware and software technologies.

1.2.5.4.2 Vision and expected outcome

For increased sustainability of ECS, the circular economy, with its 9R framework (Refuse, rethink, reduce, reuse, repair, refurbish, remanufacture, repurpose, recycle and recover) and eco-design, are the main tools to reduce the environmental footprint of ECS. Life cycle assessment as a framework will be used for identifying hotspots and checking results of the intended reduction. But ECS themselves can also be regarded as an enabler for a more circular economy.

Future ECS products must be environmentally friendly, covering all aspects from materials, manufacturing, operation and maintenance during their lifetime, considering recycling at their end-of-life. Activities must start by ensuring circularity (eco-design, environmentally friendly materials and manufacturing), employing a low CO_2 footprint over the whole life cycle, and facilitating the transition to a circular economy, wherever possible. Outcome activities will address:

- Non-fossil, recyclable, biodegradable and compostable materials, without releasing any dangerous materials or having other negative impact on the environment.
- Eco-design including eco-reliability of sustainable and more modular ECS.
- Sustainability and reducing energy consumption and environmental footprint of the manufacturing and integration processes.
- Increasing energy efficiency of ECS, during manufacturing and lifetime, and end-of-life.
- Upstream considerations and design for repair, upgradeability, dismantling, materials separation and recycling, lifetime extension and system health monitoring, self-monitoring and healing.
- Performance and condition monitoring, traceability, (predictive) maintenance, repair, upgrading, reconfiguring, recharging, retrofitting and re-use in second life, including ecosystems and tools to support these actions.
- Product indexes and digital product passports for ECS.
- Enablement of novel sustainable business models.

The development of integration processes based on new design tools will allow the dismantling of components, as their recycling and recovery of materials (urban mining) is essential. Therefore, system design techniques move towards eco-design where we need to rethink the use of multifunctional components and modules. Design for component separation and recyclability is generally required in the selection of materials and the integration technologies. Recycling technologies, as well as new approaches to second life of ECS and re-use in new applications, must be advanced. For example, with the electrification of cars, the recycling and re-use of battery packs, modules, and individual cells, and finally, materials recovery from the cells, becomes more and more important. Finally, we should of course minimize the number of batteries we use and dispose and improve recycling of batteries and battery materials.

The use of new environmentally friendly, recyclable and non-fossil materials (or compostable/biodegradable materials) must be seriously considered to replace existing materials with low recyclability in the near future. The use of these materials can easily be extended to other parts of the system, and the development of biodegradable materials can also contribute to solving the problems of recyclability. Life cycle assessment (LCA) should be used as a design tool to minimize the ECS carbon and environmental footprint, considering the full life cycle and end-of-life.

1.2.5.4.3 Key focus areas

- Eco-Design of ECS to promote circularity:
    - Use of replacement materials to comply with Restriction of Hazardous Substances Directive (ROHS) regulations (such as lead, mercury and other metals, flame retardants and certain phthalates, PFAS) and minimization of critical raw materials (CRM) dependence, including rare earths replacement for magnetics, inductors and power integrity.
    - Use of recyclable, biodegradable, compostable, non-fossil materials from sustainable sources in combination with the development of efficient and environmentally benign recycling techniques in accordance with the legislative agenda.
    - Less materials for higher functionalization.
    - Assessment of the environmental impact of ECS at the design stage, as a tool for sustainable ECS, using life cycle assessment (LCA) or similar framework.
    - Certified up-to-date data for LCA, PEF, PCR (product category rules) and EPD (environmental product declarations)
    - Development of design, fabrication, integration, recovery, reconfiguration/reuse/ repairability and disassembly strategies for products (module dismantling, component recycling, material recovery) and also short-lifetime devices (e.g. single-use medical devices, radio-frequency identification, RFID, tags and printed sensors) to meet the existing and emerging regulatory requirements.
    - Improvement of system reliability as a means to guarantee and extend the lifetime of electronic products with the final objective of responding to material efficiency requirements and providing an optimal balance on a life cycle scale.
    - Increasing power efficiency of ECS during lifetime by using low-power techniques with context awareness or energy harvesting.
    - Extensive use of software to increase the sustainability of ECS by extending product lifetime through continuous optimization and adaptation, by making existing ECS more intelligent through the use of AI, in particular at the Edge, and by optimizing the resource usage through hardware and software co-defined strategies. Ensure the trustworthiness and reliability of the ECS, including its software components, with a special attention to approaches involving AI that must be secure, reliable and automatically and autonomously adapted.
    - Promote methodologies allowing for the co-design of ECS hardware and software involving simulations and realistic models, including AI-aided development tools, to continuously estimate key metrics. This must be complemented by instrumentation of the ECS hardware and software to continuously assess the achievements of the key figures of merit over the entire ECS lifetime (including shelf and post-decommissioning).

- Sustainable manufacturing of ECS:
    - Explore the extension of materials efficient wet/dry processes such as additive manufacturing methods and laser/mechanical subtractive processes. E.g. printing, that consume less resources (energy, materials, water) and are compatible with renewable materials, such as bio-based substrates. At the same time, additive manufacturing offers new design capabilities for circular, thin and flexible devices, even for single use (e.g. wearable electrodes) with specified end-of-life management.
    - Optimization of resources and processes in production environments with potential in-situ re-use and regeneration of base materials and chemicals.
    - Reduce energy consumption (and greenhouse emissions) in clean rooms through the use of renewable energy sources and energy-efficient technologies or tools.
    - Increase water reusage in electronics manufacturing facilities.
    - Improved analytics for hazardous production remnants

    - Inline sensors to research and mitigate environmental risks of GHGs, PFAS and hazardous materials and their integration to abatement systems and process equipment
    - Improve gas abatement systems.
    - Solutions to value scarce material use and to work on alternative technology integration strategies.
    - Breakthroughs and development in recycling processes and solutions for energy storage components, such as batteries.
    - Moving towards CO_2-neutrality and zero waste ECS economy.

- Sustainable products and business models:
    - Encouraging sustainable supply chains.
    - Introduction of product category rules and product indexes (including info such as energy and resource efficiency, durability, reusability, upgradability and repairability, presence of substances that inhibit circularity, recycled content, remanufacturing and recycling, carbon and environmental footprints, expected waste generation and information requirements) for components and systems to encourage the use of LCA based environmental product declarations.
    - Digital product passports should be promoted, tested and then widely established.
    - Encourage new business models to see value in eco-design and recyclability.
    - Value repairability: An EU-wide repair index inspired by the French repairability index.
    - Condition monitoring for usage as well as for health/performance and anomaly detection.
    - Improve efficiency of e-waste recyclability by robotics, thereby increasing new value streams and business through reuse.
    - Life cycle traceability of components and systems to capture condition, carbon footprint, authenticity and recyclability.

1.2.4 Timeline

The following tables illustrate the roadmaps for **Components, Modules and Systems Integration**.

Major Challenge	Topic	Short Term (2025-2029)	Mid-Term (2030-2034)	Long Term (2035 and beyond)
Major Challenge 1: **Functionality**	**Topic 1.1:** **Sensing, imaging and actuation**	• Selective gas-sensing • Disease monitoring and diagnostics platforms (in vitro, wearables) • Lidar and radar systems • Functional materials (piezo, ceramics, polymers, metamaterials) • IR sensors integrated with CMOS • H2 low and mid pressure sensors; H2 detection in exhaust	• Selective detection of allergens, residues • Fluidics • Drug delivery • Affordable IR imagers • Hyperspectral imaging • Materials and concepts for Quantum sensors • H2 detection in ambient surrounding	• Convergence of sensing principles (e.g. thermal, optical cameras with lidar/radar) • Multifunctional healthcare support systems (wearables, implants) • Integrated Quantum sensors
	Topic 1.2: **MEMS technology**	• Novel piezo materials and piezo devices for MEMS/NEMS • Micro-optical (MOEMS) components • Compact audio MEMS • Self-monitoring and -calibration	• Integration for multifunctional sensors and actuators based on MEMS/NEMS and MOEMS • Strong mechanical actuation mechanisms • Low-power audio MEMS systems for ANC	• Self-monitoring, correcting and -adapting MEMS/NEMS • Highly integrated multifunctional, dynamically adaptive and context recognizing sensors
	Topic 1.3: **Integrated photonics**	• Novel devices operating at different wavelengths than used for telecom • Co-packaging and integration of Integrated photonics and high-speed electronics • Photonic health and medical sensors	• Tunable laser sources for PICs • Materials and devices for Quantum PICs. • Optical elements for beam shaping and manipulation (like ultra-thin curved waveguides, metalenses, tunable lenses and filters, next generation holograms, ultra-wide-angle holograms) • Display technologies (like microLEDs, MEMS-mirrors, Phase Arrays) and sensors (e.g. for eye tracking)	• Growth of light-emitting structures on silicon and integration into photonic platforms • Analogue and Neuromorphic photonic computing

	Topic 1.4: **Flexible electronics**	• Si devices compatible with integration to flexible devices; thinned IC etc. • New flexible non-fossil materials for flexible and structural electronics including active components, transparent conductors, barriers.	• Large area flexible and stretchable sensors and actuators • Organic and biocompatible materials • Wearable smart systems combing simultaneous biochemical and biophysical sensing	• Stretchable smart systems for wearables combing simultaneous biochemical and biophysical sensing • Metamaterial sensors
	Topic 1.5: **Communications**	• Real-time, low-latency, low-power, context aware, fault-tolerant and self-repairing networks for edge and IoT devices • High-speed photonics communications modules beyond 1Tb/s • Reduction of EMI	• Quantum key distribution • Advanced interconnect photonics at component as well as at system-level • Beyond 5G and 6G communications, including non-terrestrial networks • THz communication • Energy constraint aware and adaptive networks at node and network level • Accurate and stable clocks for 6G and quantum devices	• Quantum internet and cryptography • Beyond 6G • Digital twins at node and network level to help design and optimize energy constraint aware WSN architectures at planning and operational stages
	Topic 1.6: **Energy and thermal management**	• Lightweight energy harvesters and storage • Multi source energy harvesting PMIC operating down below 10mV and 10µW • Low power components • Energy storage devices for extreme temperature and harsh environments • Energy autonomous systems • Thermal management at different integration levels including advanced and active cooling systems • Multi-modal device and system level energy harvesting/power consumption simulation models	• Low/zero power components and systems • Solution for thermal management in integrated photonics • Advanced encapsulation materials for energy harvesters • Extend chiplet concept (design and manufacturing) to no-IC components • Sensors and actuators for the optimization of battery cells usage during their entire lifetime	• CO2-neutrality and circular economy for ECS • Energy harvesting PMICs embedded in MEMS & NEMS WSN nodes with MCU, sensors, transceivers, etc.. • Sensors embedded in energy source components for performance and condition monitoring, lifetime provenance and anomaly detection

	Topic 1.7: **Information processing**	• Security and privacy • Explainable AI, edge computing (HW and SW) • Hybrid modelling (physical and data-driven) • Federated data collection from edge to gateway to cloud, to minimize strain • Energy constraint aware and adaptive networks and architectures, particularly for battery-powered edge devices	• Integration of information processing close to data acquisition • Hardware solutions for security and privacy • Neuromorphic computing • AI in the edge computing • Quantum simulation and quantum computing for the data-analysis (in the cloud)	• Low-power AI • Neuromorphic on-the-edge computing for sensors and actors • Quantum computing • Quantum simulation

Major Challenge 2: **Advanced integration solutions**	**Topic 2.1:** **Advanced packaging and SiP technologies**	• Integration for complexity: Hybrid integration of heterogeneous components into several types of platforms • System health monitoring and self-diagnosis • Integration of biological and molecular functions, integration with fluidics • Embedding of power sources (batteries, energy harvesting) in SiP and IC carriers with digital interfacing • Embedding of power sources, ASIC drivers and audio MEMS sensor and actuators such as microphones and micro speakers into a heterogeneous SiP	• Integration for harsh environments, and implantable electronics • System health monitoring and self-diagnosis, self-healing • Self-cleaning and self-healing materials • Integration of different sensors and sensor hubs for sensor fusion (e.g. combination of acceleration sensor, microphone, microspeaker for enhanced noise cancellation).	• Maximum functional integration in minimum volume/footprint Advanced photonics • Biological-electronics hybrid systems
	Topic 2.2: **Integrated Photonics and co-integration with electronics**	• Photonics integration with RF, sensors; electro-optic co-packaging • High-precision component placement and bonding processes and equipment • Low-loss fiber coupling to PICs.	• Heterogeneous integration of active components (e.g. III-V) on PICs on wafer scale. • Metamaterials for beam shaping	• Combining electrical and optical interconnects into an electro-optical IC carrier • Monolithically integrated quantum photonics including III-V quantum dots

Major Challenge 3: **Heterogeneous integration**	**Topic 3.1:** **Flexible electronics**	• Integration processes with flexible, structural and 3D conformable electronics • Materials for chip interconnection; ACA, ICA, flip chip etc. • Large area R2R compatible Interconnection processes and equipment for heterogeneous integration	• Compostable and biodegradable substrate and housing materials • Roll-to-Roll compatible chip assembly and interconnection technologies on stretchable substrates	• Stretchable electronics and system integration • Automated Interconnection processes and equipment for heterogeneous integration
	Topic 3.2: **Quantum systems**	• Materials and methods for quantum technology integration in cryogenic temperatures • Materials and components for low-loss and high quantum coherence • Cryogenic electronics, cryo-CMOS	• Integration of quantum systems: superconducting, photonic, Silicon technologies • Solid state coolers	• Quantum SiP also in room temperature or with integrated cooling.
	Topic 3.3: **Manufacturing methodology, characterization and testing**	• I4.0 for manufacturing optimization • Additive manufacturing and rapid prototyping technologies and materials • Improved automation and customization in integration for smaller lots • Database of material properties for simulation and reliability • process model implementation as Digital Twins for Manufacturing Optimization	• I4.0 for manufacturing optimization, Zero-defect integration • Automation and customization in integration for smaller lots. • Additive manufacturing and rapid prototyping technologies • Material by design approach	• Automation and customization in integration, lot one. • Fully digitalized manufacturing process description – Digital Twin for Manufacturing

Major Challenge 4: **Sustainability**	**Topic 4.1:** **Eco-design of ECS**	• Replacement materials to comply with RoHS (e.g. PFAS) and minimize CRM dependence • Use of recyclable materials (e.g. moulds, resins) • Less materials for higher functionality • Life cycle analysis as tool for design • Development of design, fabrication, integration, recovery, reconfiguration/reuse, disassembly strategies • Designing for repairability, including modular approach, upgrades and maintenance • Certified up-to-date data for LCA, PEF, PCR and EPD • Eco-design benchmark values for electronic components • Availability and exchangeability of spare parts and tools • Improve system reliability to guarantee and extend the lifetime • Use low-power techniques with context awareness and/or energy harvesting	• Cross-company reuse of "stable" chip designs, including More-than-Moore components • Methodologies allowing for the co-design of ECS hardware and software • Use of biodegradable, compostable, non-fossil materials • Breakthroughs in recycling processes, including energy storage components • Applying results of Green ECS to integrated photonics • Extensive use of SW to increase sustainability of ECS • Holistic sustainability assessment (social, economic, environmental) as tool for design • Recycling and materials for recycling (system level) • Energy efficient processes for manufacturing • Build sharing of information on LCA and Safe and sustainable by Design	• Highly integrated re-usable circuit blocs • Circular economy of ECS • Solutions for full recycling and material recovery of ECS, including energy storage components • Environmental footprint and critically based recycling planning (final stage material recycling) • Set up repair process: failure characterization, repair and re-characterization; provide manuals, instructions, schematics and inexpensive spare parts • Warranties and safety-relevance need to be addressed

	Topic 4.2: **Sustainable manufacturing of ECS**	• Condition monitoring and predictive maintenance • Increasing the energy and resource efficiency and environmental footprint of components, systems and modules • Use of energy harvesting to minimize/eliminate battery replacement • Extension of additive manufacturing methods • Increase water reusage • Improved analytics for hazardous production remnants	• Reducing energy consumption and CO_2 footprint of integration processes, tools and ECS systems • Inline sensors to research and mitigate environmental risks of GHGs, PFAS and hazardous materials and their integration to abatement systems and process equipment • Processes for reuse and second life • Improved recycling and material recovery processes • Replacement of critical process steps to comply with RoHS (e.g. PFAS) • Work on alternative technology integration strategies	• CO_2-neutral ECS economy • Zero waste added manufacturing to produce functional modules • Recycling processes adapted for flexible, sustainable electronics with novel materials
	Topic 4.3: **Sustainable products and business models**	• Establish a repair index inspired by the French repairability index • Introduce product category rules and product indexes	• Self- condition monitoring energy harvesting solutions to assure long term reliability and for anomaly detection • Repair as business (bonus-malus-systems) • A business model can emerge for third-party repair centers • Closing data gaps in circularity and recycling through Digital Product Passport • Improve efficiency of e-waste recyclability by robotics • Life cycle traceability of components and systems to capture condition, carbon footprint, authenticity and recyclability	• Train skilled repairers; questions around re-certification

1.3

Foundational Technology Layers

EMBEDDED SOFTWARE AND BEYOND

1.3.1. Introduction

The Artemis/Advancy report[1] states that "*the investments in software technologies should be on at least an equal footing with hardware technologies, considering the expected growth at the higher level of the value chain (Systems of Systems, applications and solutions)*". According to the same report, embedded software and software engineering tools are part of the six technology domains needed for embedded intelligence. Embedded intelligence means incorporating AI algorithms ("classic" or ML ones) in devices or components to give them the ability to reflect on their own state (e.g. operational performance, usage load, environment), execute tasks independently, adjust to novel circumstances, and make data-driven decisions without human input. Such devices will operate in a robust and resilient way, e.g. independent of internet connectivity and are the necessary step towards the next level of digitalisation and sustainability. In this context, embedded intelligence supports the green deal initiative, as one of the tools for enhancing sustainability.

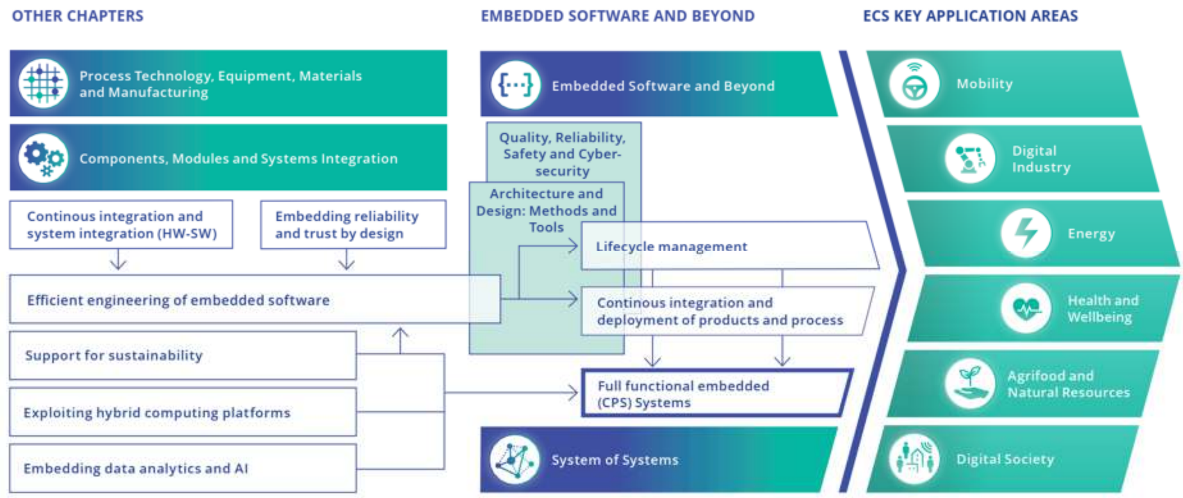


Figure 1.3.1 - Positioning of the Embedded Software and Beyond Chapter in the ECS-SRIA

Figure 1.3.1 illustrates the role and positioning of the **Embedded Software and Beyond** Chapter in the ECS- SRIA. The Chapter on **Components Modules and System Integration** focuses on functional hardware components and systems that compose the embedded and cyber-physical systems (CPS), considered in this Chapter. While the **System of Systems** (SoS) Chapter is based on independent, fully functional systems, products and services (which are also discussed in this Chapter), they are also the constituents of SoS-based solutions. The **Architecture and Design: Methods and Tools** Chapter examines engineering processes, methods, and tools, while this Chapter focuses more on the technology stack of **Embedded Software and Beyond**. For the discussion on safe, trustworthy, and explainable AI in the context of embedded intelligence, this Chapter is also linked to **Quality, Reliability, Safety and Cybersecurity (**Chapter 2.4).

[1] Advancy, 2019: Embedded Intelligence: Trends and Challenges, A study by Advancy, commissioned by ARTEMIS Industry Association. March 2019. Downloadable from: https://www.inside-association.eu/publications

This Chapter is called **Embedded Software and Beyond** to stress that embedded software is more than "just software": it is a key component of any system's embedded intelligence, it enables systems to act on external events, and it enables inter-system communication.

Most importantly, Embedded software empowers Embedded and cyber-physical systems (ECPS) to play a key role in solutions for digitalisation in almost every application domain (cf. Chapters 3.1-3.6). From a functional perspective, the role of Embedded Software is becoming increasingly dominant because of the new software-enabled functionalities ECPS (e.g. cars, trains, airplanes and health equipment) need to provide (including aspects as security, privacy and autonomy). In these systems, most of the innovation comes from software, nowadays. ECPS also form the backbone of SoS (e.g. smart cities, air traffic management), providing required interconnection and interoperability. Owing to all these factors, ECPS are an irreplaceable part of the strive towards digitalisation of our society.

At the same time, ECPS need to exhibit required quality properties (e.g. safety, security, reliability, dependability, sustainability, and, ultimately, trustworthiness). Furthermore, due to their close integration with the physical world, ECPS must consider the dynamic and evolving aspects of their environment to provide deterministic, high-performance, and low-power computing, especially when processing intelligent algorithms. Increasingly, software applications will run as services on distributed SoS involving heterogeneous devices (e.g: servers, edge devices) and networks, with a diversity of resource restrictions. In addition, it is required from ECPS that their functionalities and hardware capabilities evolve and adapt during their lifecycles – e.g. through updates of software or hardware in the field and/or by learning. Building these systems and guaranteeing their previously mentioned quality properties, along with supporting their long lifetime and certification, requires innovative technologies in the areas of modelling, software engineering, model-based design, verification and validation (V&V) technologies, and virtual engineering. These advances need to enable engineering of high-quality, certifiable ECPS that can be produced (cost-)effectively (cf. Chapter 2.3, **Architecture and Design: Methods and Tools**).

1.3.2. Scope

Common challenges in embedded software and its engineering for ECPS include:

- Interoperability.
- Complexity of requirements and code (safety, security, performance).
- Quality (dependability, sustainability, performance, trustworthiness).
- Lifecycle (maintainability, extendibility).
- Efficiency, effectiveness, and sustainability of software development.
- Adaptability to, and the dynamic environment of ECPS.
- Maintenance, integration, rejuvenation of legacy software solutions.

To enable ECPS functionalities and their required level of interoperability, the engineering process will be progressively automated and will need to be integrated in advanced SoS engineering covering the whole product during its lifetime. Besides enabling new functionalities and their interoperability, it will need to cover non-functional requirements (safety, security, run-time performance, reliability, dependability, sustainability, and, ultimately, trustworthiness) visible to end users of ECPS, and to also satisfy quality requirements important to engineers of the systems (e.g. evolution, maintenance). This requires innovative technologies that can be adapted to the specific requirements of ECPS and, subsequently, SoS.

Further complexity will be imposed by the introduction of Artificial Intelligence (AI), machine-to-machine (M2M) interaction, new business models, and monetisation at the edge. This provides opportunities for enhancing new engineering techniques like AI for SW engineering, and SW engineering for AI. Future software solutions in ECPS will solely depend on new software engineering tools and engineering processes (e.g. quality assurance, Verification and Validation (V&V) techniques and methods on all levels of individual IoT and in the SoS domain).

Producing industrial software, and embedded software in particular, is not merely a matter of writing code: to be of sufficient quality, it also requires a strong scientific foundation to assure correct behaviour under all circumstances. Modern software used in products such as cars, airplanes, robots, banks, healthcare systems, and the public services comprises millions of lines of code. To produce this type of software, many challenges have to be overcome. Even though software in ECPS impacts everyone everywhere, the effort required to make it reliable, maintainable and usable for longer periods is routinely underestimated. As a result, every day there are news articles about expensive software bugs and over-budget or failed software development projects. Also, big challenges with correctness and quality properties of software exist, as human well-being, economic prosperity, and the environment depend on it. There is a need to guarantee that software is maintainable and usable for decades to come, and there is a need to construct it efficiently, effectively and sustainably. Difficulties further increase when legacy systems are considered: information and communications technology (ICT) systems contain crucial legacy components at least 30 years old, which makes maintenance difficult, expensive, and sometimes even impossible.

The scope of this Chapter is research that facilitates engineering of embedded software used for ECPS, enabling digitalisation through the feasible and economically accountable building of SoS with necessary quality. It considers:

- Challenges that arise as new applications of ECPS emerge.
- Continuous integration, delivery and deployment of products and processes.
- Engineering and management of ECPS during their entire lifecycle, including sustainability requirements.

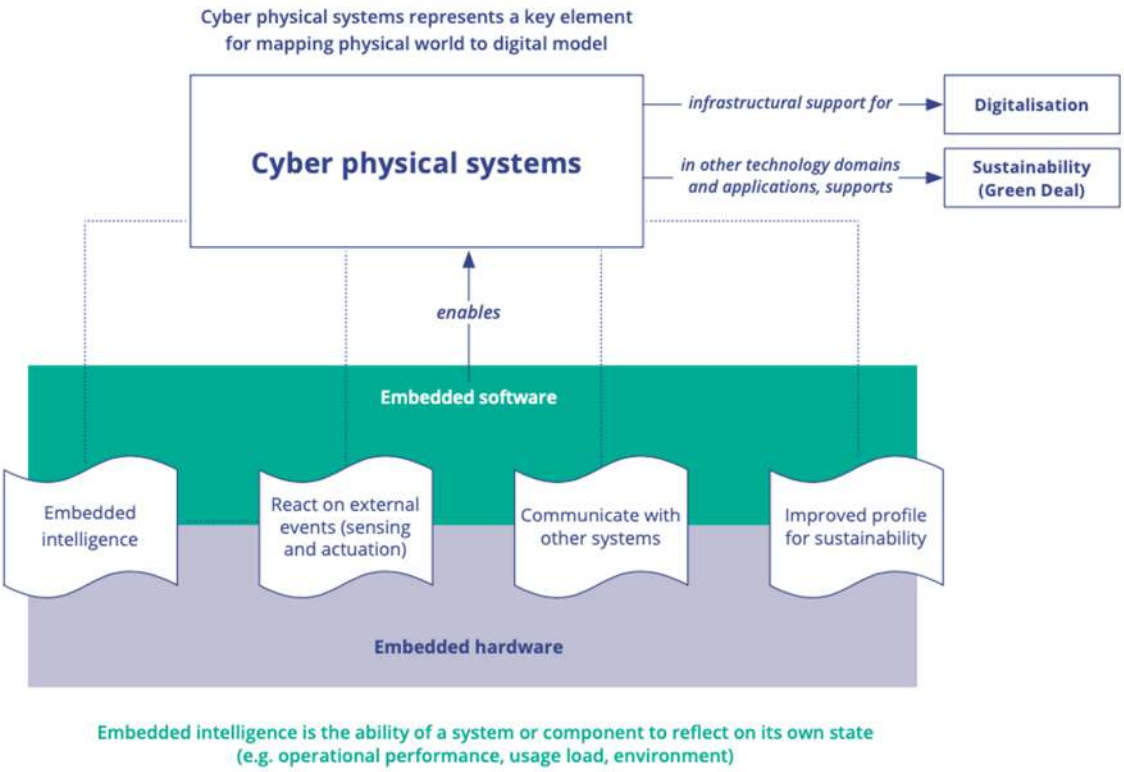


Figure 1.3.2 – Importance of Embedded Software for Cyber physical systems and its roles.

Quantum Technologies

Quantum technology has drawn a growing amount of attention in recent years. This short text briefly explains the three main topics of this fields. An inventory is made of the impact of quantum technologies on embedded software and beyond.

The Three Main Topics of Quantum Technology
Quantum computing, quantum internet and quantum sensing are the three main topics of quantum technology. Let's take a look at all three.

Quantum Computing Quantum computing is the most captivating of the three. In theory, quantum computers are able to solve some types of computations exponentially faster than classical computers. Shor's algorithm to factor a number in its prime factors is often quoted as an example of this speedup. This opens the road to find cryptographic keys which form the backbone of today's secure communication technologies. Once quantum computers become practical, this may well pose a threat to secure, encrypted communication. The key point here is that this threat requires a significantly larger quantum computer than is available now: for factoring a key of several thousand bits, a logical qubit register of several thousand qubits is required. With the current state-of-the-art, this requires millions of physical (noisy) qubits. Such a large quantum computer is at least 10 to 15 years away, if not more[2].

[2] Preskill, J. (2018). Quantum Computing in the NISQ era and beyond. Quantum, 79.

Nevertheless, quantum computing is casting its shadow ahead. And governments and organisations are already taking countermeasures. Several governments now require all of their services to prepare for the security threat quantum computing will pose. It means that current encryption technologies are to be upgraded to a degree that even challenges quantum computing. This asks for longer encryption keys and more complex encrypting and decrypting algorithms, requiring more resources. Research and development of efficient digital cryptography systems, involving hardware and software, is already ongoing and will play an increasingly important role as quantum computers are coming of age[3].

These observations lead to the conclusion that quantum computing will have an indirect impact on embedded software and beyond, in the next few years. It depends on the speed of evolution and innovation of quantum technology when quantum computing devices will leave the laboratory and make their introduction to the industry. For now, that appears to be a decade away, but vigilance on this subject is required. Europe should strive for independence from other nations in this area to be able to develop this technology on its own, in the light of the recent developments in international relations.

Quantum Internet

A quantum internet is an application of quantum networks. Quantum networks enable the communication of qubits. Such networks can be used to connect quantum processors to form more powerful quantum computers. Quantum networks can also be used to create quantum internet applications. One such application is the secure distribution of cryptographic keys: in this setup, cryptographic keys are distributed over a quantum network using entangled qubits, enabling the detection of eavesdropping on the communication. But quantum internet, just like quantum computers, are still in the research and development phase. Practical applications at this moment require complicated setups, often involving cryogenically-cooled devices, preventing wide-spread use today and in the next few years[4].

Quantum Sensors

Quantum sensors are sensors which detect physical properties by using quantum effects such as quantum entanglement, quantum interference, and quantum state squeezing. Quantum sensors have been in use for quite a long time: medical magnetic resonance scanners, which detect the precession of atomic nuclei in a magnetic field.

Quantum sensors are sensitive to some physical property. It is not so much the measurement of the physical property, but the enhanced accuracy or sensitivity to such a property that makes quantum sensors stand out from classical sensors. As such, the (embedded) software that processes the measurement of quantum sensors does not differ from software that processes measurements from classical sensors[5].

---

[3] Post-Quantum Cryptography - Setting the Future Security Standards. (n.d.). Retrieved from https://www.nxp.com/applications/enabling-technologies/security/post-quantum-cryptography:POST-QUANTUM-CRYPTOGRAPHY

[4] Singh, A., Dev, K., Siljak, H., Joshi, H., & Magarini, M. (2021). Quantum Internet—Applications, Functionalities, Enabling Technologies, Challenges, and Research Directions. IEEE Communications Surveys & Tutorials, 2218-2247. doi:10.1109/COMST.2021.3109944

[5] Kantsepolsky, B., Aviv, I., Weitzfeld, R., & Bordo, E. (2023). Exploring Quantum Sensing Potential for Systems Applications. IEEE Access, 31569-31582.

It appears that quantum technology will impact the communication security of embedded systems in the next few years. Implementations for post-quantum cryptography must be researched and developed to stay ahead of quantum technology developments.

1.3.3. APPLICATION BREAKTHROUGHS

Embedded software significantly improves the functionalities, features, and capabilities of ECPS, increasing their autonomy and efficiency, and exploiting their resources and computational power, as well as bringing to the field functionalities that used to be reserved only for data centres, or more powerful and resource-rich computing systems. Moreover, implementing specific functionalities in software allows for their re-use in different embedded applications due to software portability across different hardware platforms. Examples of increasing computational power of ECPS are video conferencing solutions: less than 20 years ago specialised hardware was still required to realise this function, with big screens in a dedicated set-up that could not be used for any other but a dedicated application. Today, video conferencing is available on every laptop and mobile phone, where the main functionality is implemented by software running on standard hardware. The evolution is pushing to the “edge” specific video conferencing functionalities, adopting dedicated and miniaturised hardware supported by embedded software (video, microphone, and speakers), thus allowing the ECS value chain to acquire a new business opportunity.

Following a similar approach, it has been possible to extend the functionalities of mobile phones and smart watches, which today can a.o. count steps, keep track of walked routes, monitor health, inform users about nearby restaurants, all based on a few extra hardware sensors and a myriad of embedded software applications. The trend is to replace specialised hardware applications with software running on generic computing hardware and supported by application-specific hardware, such as AI accelerators, neural chips. This trend is also contributing to the differentiation of the value creation downstream and upstream, as observed in the Advancy report [6] (see Figure 1.3.3).

These innovations require the following breakthroughs in the field of embedded software:

- Increased engineering efficiency and an effective product innovation process (cf. Chapter 2.3 Architecture and Design: Methods and Tools).
- Enabled adaptable systems by adaptable embedded software and machine reasoning.
- Improved system integration and verification and validation.
- Embedded software, and embedded data analytics and AI, to enable system health monitoring, diagnostics, preventive maintenance, and sustainability.
- Data privacy and data integrity.

[6] Advancy, 2019: Embedded Intelligence: Trends and Challenges, A study by Advancy, commissioned by ARTEMIS Industry Association. March 2019. Downloadable from: https://www.inside-association.eu/publications

- Model-based embedded software engineering and design as the basis for managing complexity in SoS (for the latter, cf. Chapter 2.3 Architecture and Design: Methods and Tools).
- Improved multidisciplinary embedded software engineering and software: architecting/design for (systems) qualities, including reliability, trust, safety, security, overall system performance, installability, diagnosability, sustainability, and re-usability (for the latter, cf. Chapter 2.3 Architecture and Design: Methods and Tools and Chapter 2.4 Quality, reliability, safety and cybersecurity).
- Upgradability, dealing with variability, extending lifecycle and sustainable operation.

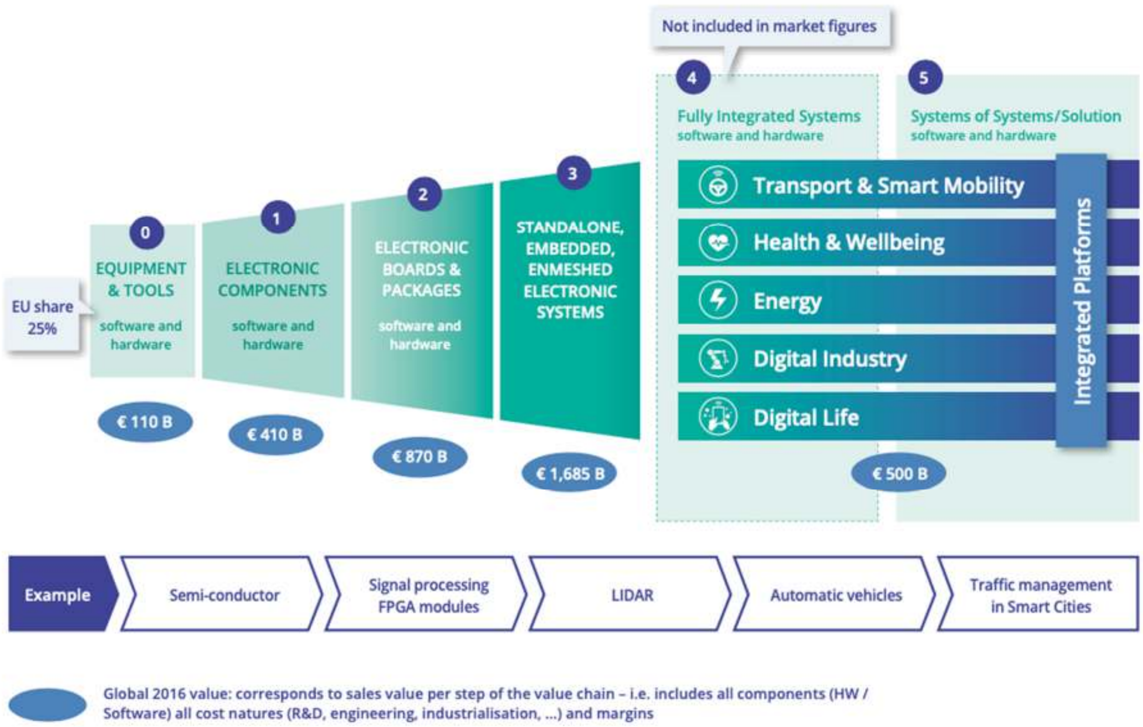


Figure 1.3.3 - Advancy (2019) [7]report: value creation

[7]Advancy, 2019: Embedded Intelligence: Trends and Challenges, A study by Advancy, commissioned by ARTEMIS Industry Association. March 2019. Downloadable from: https://www.inside-association.eu/publications

1.3.4. MAJOR CHALLENGES

Research and innovation in the domain of embedded software and beyond will have to face seven challenges, each generated by the necessity for engineering automation across the entire lifecycle of sustainability, embedded intelligence and trust in embedded software.

- **Major Challenge 1**: Efficient engineering of embedded software.
- **Major Challenge 2**: Continuous integration and deployment.
- **Major Challenge 3**: Lifecycle management.
- **Major Challenge 4**: Embedding data analytics and artificial intelligence.
- **Major Challenge 5**: Support for sustainability by embedded software.
- **Major Challenge 6**: Software reliability and trust.
- **Major Challenge 7:** Hardware virtualization for efficient SW engineering.

1.3.6.1 Major Challenge 1: Efficient engineering of embedded software

1.3.6.1.1. State of the art

Embedded software engineering is frequently more a craft than an engineering discipline, which results in inefficient ways of developing embedded software. This is visible, for instance, in the time required for the integration, verification, validation and release of embedded software, which is estimated to exceed 50% of the total R&D&I expenses.

A new set of challenges to engineering embedded software is introduced with the emergence of heterogeneous computing architectures into the mainstream. It will be common for embedded systems to combine several types of accelerators to meet power consumption, performance requirements, safety, and real-time requirements. Development, optimisation, and deployment of software for these computing architectures proves to be challenging. If no solutions are introduced which automatically tailor software to specific accelerators[8, 9], developers will be overwhelmed with the required effort.

Software engineering is exceeding the human scale, meaning it can no longer be overseen by a human without supporting tools, in terms of velocity of evolution, and the volume of software to be designed, developed and maintained, as well as its variety and uncertainty of context. Engineers require methods and tools to work smarter, not harder, and need engineering process automation and tools and methods for continuous lifecycle support. To achieve these objectives, we need to address the following practical research challenges:

---

[8]Advancy, 2019: Embedded Intelligence: Trends and Challenges, A study by Advancy, commissioned by ARTEMIS Industry Association. March 2019. Downloadable from: https://www.inside-association.eu/publications

[9]https://www.intel.com/content/www/us/en/developer/articles/technical/efficient-heterogenous-parallel-programming-openmp.html#gs.85zv3a

shorter development feedback loops; improved tool-supported software development; methods and tools to enable strongly linked, yet independent and heterogenous development processes in new areas like software defined vehicles (SDV); empirical and automated software engineering; and safe, secure and dependable software platform ecosystems.

1.3.6.1.2. Vision and expected outcome

The demand of embedded software is higher than we can humanly address and deliver, exceeding human scale in terms of evolution speed, volume and variety, as well as in managing complexity. The field of embedded software engineering needs to mature and evolve to address these challenges and satisfy market requirements. In this regard, the following four key aspects must be considered.

(A) From embedded software **engineering to cyber physical systems engineering**

Developing any high-tech system is, by its very nature, a multi-disciplinary project. There is a whole ecosystem of models (e.g. physical, mechanical, structural, (embedded) software and behavioural) describing various aspects of a system. While many innovations have been achieved in each of the disciplines separately, the entirety still works in silos, each with their own models and tools, and only interfacing at the borders between them. This traditional separation between the hardware and software worlds, and individual disciplines, is hampering the development of new products and services.

Instead of focusing only on the efficiency of embedded software engineering, we already see that the field is evolving into direction of cyber physical systems (cf. Chapter 2.3 **Architecture and Design: Methods and Tools**), and software is one element of engineering. Rather than silos and handovers at the discipline's borders, we expect tools to support the integration of different engineering artefacts and enable, by default, effective development with quality requirements in mind – such as safety, security, reliability, dependability, sustainability, trustworthiness, and interoperability. New methods and tools will need to be developed to further facilitate software interaction with other elements in a system engineering context (cf. Chapter 2.3 **Architecture and Design: Methods and Tools**).

Software-defined systems enable the implementation of complex and customisable functionalities in CPS. The equivalent for the automotive industry is the software-defined vehicle (SDV), which often includes the concept of the connected vehicle and the associated cloud services. SDVs combine mechatronic - often safety-critical and real-time-capable - systems with edge, internet, app and cloud technologies. This requires the integration of embedded software (open and closed source), which in turn is often developed according to different paradigms, standards and business models. In the target vehicle system, these parts must not only function together, but also fulfill strict quality standards and, where applicable, legal regulations. SOA-based approaches are expected to contribute significantly towards these goals. Methods and tools are required that support individual but linked development cycles in the respective paradigms and at the same time enable the composition of the overall system, but also the decomposition of results from diagnostics, verification and validation back to the individual part.

Artificial intelligence is a technology that holds a great potential in dealing with large amounts of data, and potentially could be used for understanding complex systems. In this context, artificial intelligence has the potential to automate some daily engineering tasks, moving boundaries of type and size of tasks that are humanly possible in software engineering.

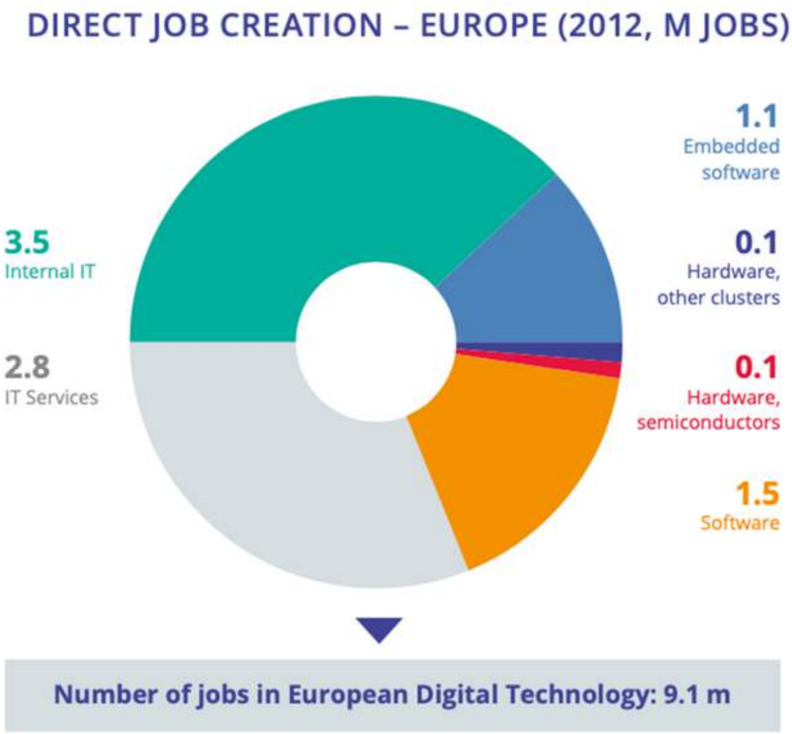


Figure 1.3.4 - Direct job creation – Europe (2012, m jobs) Source: EU, IDC, Destatis, Roland Berger

(B) Software architectures for optimal edge computing

At the moment, Edge computing lacks proper definition and, including many different types of managed and unmanaged devices, this leads to uncertainty and difficulties on how to efficiently and effectively use software architectures, including aspects as resource, device, and network management (between edge devices as well between edge and fog/cloud), security, useful abstractions, privacy, security, reliability, and scalability. Additionally, automatic reconfiguration, adaptation and re-use face a number of challenges. These challenges are caused by diversity of edge devices and wide range of requirements in terms of Quality of Service (e.g. low latency, high throughput). In addition, sustainability and reliability are difficult to ensure when trying to prioritize between Quality of Service on the edge and end-to-end system Quality of Service.
Furthermore, the lack of definition also hampers the growing need for energy efficient computing and the development of energy consumption solutions and models across all layers from materials, via software architecture to embedded/application software. Energy efficiency is vital for optimal edge computing.
Lastly, as AI is also moving towards the edge (i.e., Edge AI) defining lightweight models and model architectures that can deal with low amounts of data available on the edge and still provide good model accuracy are desperately needed. Finally, this limits transfer of common

solution patterns, best practices, and reference architectures, as Edge computing scope and configuration requires further clarification and classification.
Since edge devices need to be self-contained, edge software architectures need to support, from the one side, virtual machine-like architectures, and from the other side they need to support the entire software lifecycle. The fact that there are many different types of edge devices also requires an interoperability standard to ensure that they can work together. Innovations in this field should focus on, amongst others, software-hardware co-design, virtualisation and container technologies and new standard edge software architecture (middleware).
It is essential to discuss types of quality properties that become more significant as Edge computing is introduced, and based on these, build use cases that profit from quality properties specific to edge computing. There is a need for new approaches that enable early virtual prototyping of edge solutions, as well as approaches that enable verification and validation of quality properties during the entire life cycle of edge software systems. One of the possibilities for profiting from Edge is to focus on digital twins to monitor divergences from expected behaviour and implement logic that will benefit from Edge's low latency when making critical decisions, especially in safety critical software systems.

(C) Integration of embedded software

To ensure software development is more effective and efficient, it is necessary to place greater focus on integrating embedded software into a fully functional system. First, innovation in continuous system integration must include more effective ways of integrating legacy components into new systems (see also D). Second, for the integration of data and software, the embedded software running in the field has to generate data (such as on run-time performance monitoring, system health, quality of output, compliance to regulations, user interactions) that can be re-used to improve its quality and performance. By improving this, the data and software integration can not only improve the efficiency of embedded software itself, but also the internal coordination and orchestration between components of the system by ensuring a rapid feedback cycle. Third, it is paramount to enable closer integration of software with the available computing accelerators. This must be done in a way that frees developers from additional effort, while at the same time uses the full potential of heterogeneous computing hardware.

(D) Using abstraction and virtualisation

The recent focus on model-driven (or "low-code") software development has sparked a new approach to managing complexity and engineering software. Generating embedded software from higher-level models can improve maintainability and decrease programming errors, while also improving development speed. However, creating and managing models of real systems with an appropriate level of detail that allows for simulation and code generation is a challenge. Managing models and their variability is a necessity if we want to prevent shifting the code legacy problem to a model legacy problem where there are too many models with too much variety.

The core elements of the domain are captured in a language of the domain. The introduction of domain-specific languages (DSLs) and aspect-oriented languages has allowed for the inclusion of aspects and constructs of a target application domain into the languages used to develop embedded software. This abstraction allows for shortening the gap between software engineers and domain experts. We expect innovations in DSLs and tool support to establish a major boost in the efficiency of embedded software development.

The increased level of abstraction allows for more innovation in virtualisation of systems and is a step towards correctness by construction instead of correctness by validation/testing. Model-based engineering and digital twins of systems are already being used for a variety of goals – such as training, virtual prototyping and log-based fault analysis. Furthermore, they are necessary for supporting the transition towards sustainable ECPS. Innovations in virtualisation will allow DSLs to be (semi-)automatically used to generate digital twins with greater precision and more analysis capabilities, which can help us to explore different hardware and software options before a machine is even built, shortening development feedback loops due to such improved tool-supported software development.

(E) Resolving legacy

Legacy software and systems still constitute most of the software running in the world today. It is only natural that the amount of legacy software increases in the future. While it is paramount to develop new and improved techniques for the development and maintenance of embedded software, we cannot ignore the systems currently in operation. New software developed with novel paradigms and new tools will not run in isolation, but rather have to be used increasingly in ecosystems of connected hardware and software, including legacy systems.

There are two main areas for innovation here. First, we need to develop efficient ways of improving interoperability between new and old. With years of development, and a need to continue operations, we will have to depend on legacy software for the foreseeable future. It is therefore imperative to develop new approaches to facilitating reliable and safe interactions, including wrapping old code in re-usable containers. Second, we must innovate the process to (incrementally) migrate, rejuvenate, redevelop and redeploy legacy software, both in isolation and as part of a larger system. We expect innovations in these areas to increase efficiency and effectiveness in working with legacy software in embedded software engineering.

1.3.6.1.3. Key focus areas

The key focus areas in the domain of efficient embedded software engineering include the following. There is a strong relation between Major Challenges 1, 2, 3, and 6 below, and Chapter 2.3 "Architectures and Design: Methods and Tools“, specifically Major Challenges 1 and 2.

- Model-based software engineering:
  - Model-based software engineering enabling systems to become part of SoS.
  - Model inference to enable re-use, refactoring and evolution of existing subsystems in SoS.

    - Model-based testing that takes the re-use of uncontrolled systems into account.
    - Embedded software architectures to facilitate building SoS.

- Digital twinning:
    - Virtualisation as a means for dealing with legacy systems.
    - Virtualisation and virtual integration testing (using Digital Twins and specialized design methods, like, e.g. contract-based design, for guaranteeing safe and secure updates (cf. Architecture and Design: Methods and Tools Chapter 2.3).
    - Approaches to reduce re-release/re-certification time, e.g. model-based design, contract-based design, and modular architectures.
    - Distinct core system versus applications and services.
    - Design for X (e.g design for test, evolvability and updateability, diagnostability, and adaptability).

- Constraint environments:
    - Knowledge-based leadership in design and engineering.
    - Resource planning and scheduling (including multi-criticality, heterogeneous platforms, multicore, software portability).
    - Simulation and Design for software evolution over time, while catering for distinct phases.
    - Exploiting hybrid compute platforms, including efficient software portability.

- Software technology:
    - Virtualisation as a tool for efficient engineering.
    - Interface management enabling systems to become part of SoS.
    - Technology for safe and dependable software ecosystems.
    - Artificial intelligence-based tools to support software engineering, software production and testing efforts.
    - Co-simulation platforms.

- SW engineering tools:
    - Integrating embedded AI in software architecture and design.
    - Programming languages for developing large-scale applications for embedded systems.
    - Models & digital twins, also at run-time for maintainability and sustainability.
    - Programming models, compilers, code generators, and frameworks for optimal use of heterogeneous computing platforms.
    - Co-simulation platforms.
    - Tools, middleware and (open) hardware with permissible open-source licenses.
    - Methods and tools that support individual but linked development cycles in the respective paradigms (safety, real-time, edge/cloud, open/closed source)

and velocities on the one hand, and integration, verification and validation in complex target systems like SDV on the other hand.

1.3.6.2 Major Challenge 2: Continuous integration and deployment

1.3.6.1.4. State of the art

It is fair to assume that most future software applications will be developed to function as part of a certain platform, and not as stand-alone components. In some embedded system domains, this idea has been a reality for a decade (e.g. in the AUTomotive Open System Architecture (AUTOSAR) partnership, which was formed in 2003). Increasingly the platforms have to support SoS and IoT integration and orchestration, involving a large amount of diverse small devices. Guaranteeing quality properties of software (e.g. safety and security) is a challenging task, and one that only becomes more complex as the size and distribution of software applications grow, especially if software is not properly designed for its intended operational context (cf. Chapter 2.3 Architecture and Design: Methods and Tools). Although we are aiming towards continuous integration on the level of IoT and SoS, we are still struggling with the integration of code changes from multiple contributors into a single software system.

One aspect of the problem relates to the design of SoS[10], which are assumed to be composed of independent subsystems but over time have become dependent. Orchestration between the different subsystems, that may involve IoT as well, is an additional issue here. Another aspect relates to the certification of such systems that requires a set of standards. This applies especially for IoT and SoS and it is complicated by the introduction of AI into software systems. Although AI is a software-enabled technology, there are still many issues on the system level when it comes to its integration into software systems. It is particularly challenging to ensure their functional safety and security, and thus to certify such systems. Some of the existing initiatives include, e.g. for vehicles, ISO 21448 (Road Vehicles – Safety Of The Intended Functionality (SOTIF)), ISO/TR 4804 (followed by ISO/AWI TS 5083, currently in development), ANSI/UL 4600 Standard for Safety for the Evaluation of Autonomous Products, and SAE J3016, which recommends a taxonomy and definitions for terms related to automated driving. Note, that AI may be applied as an engineering tool to simplify certification.

Finally, integration and delivery practices are part of the engineering processes. Although methodologies already exist to achieve this (such as DevSecOps and ChatOps), these mostly relate to software production. With ECPS, continuous integration becomes increasingly more complex due to the wide range of hardware architectures and platforms, each with its own unique characteristics. Continuous integration must account for this diversity, requiring cross-compilation and testing on various target devices. Continuous delivery must deal with the fact that the products into which the new software modules have to be delivered are already sold and 'working in the field', often in many different variants (i.e. the whole car fleet of an OEM). Even in domains where the number of variant systems is small, retaining a copy of each

[10] https://www.khronos.org/sycl/

system sold at the producing company in order to have a reference target is prohibitive. Thus, virtual integration using model-based design methods (including closed-box models for legacy components) and digital twins used as integration targets as well as for verification & validation by physically accurate simulation are a mandatory asset for any system company to manage the complexity of ECPS and their quality properties. System engineering employing model-based design and digital twins must become a regular new engineering activity.

1.3.6.1.5. Vision and expected outcome

Europe is facing a great challenge with the lack of platforms that are able to adopt embedded applications developed by individual providers into an ecosystem (cf. Reference Architectures and Platforms in Chapter 2.3). The main challenges here are to ensure the adequate functionality of integrated systems (which is partially solved by the micro-services approach), while ensuring key quality properties such as performance, safety, and security (see also Major Challenge 6), which is becoming increasingly complex and neglected as we adopt approaches that facilitate only integration on the functional level. Instrumental for these challenges is the use of integration and orchestration platforms that standardise many of the concerns of the different parts in the SoS, some of which are connected via IoT. In addition, Automated engineering processes such as CI/CD will be crucial to adapt. The primary DevOps methodology needs to be adjusted for the ECPS. For example, CI pipelines will enable toolchain selection configuration for cross-compiling, or they will include real-time testing and validation, which can be more challenging to automate and verify. Integrating automated tests for hardware interactions can be complex and require specialized hardware-in-the-loop (HIL) testing setups.
ECPS will become a part of an SoS and eventually SoECPS. SoS challenges like interoperability, composability, evolvability, control, management and engineering demand ECPS to be prepared for a life as a part of a SoS (cf. Chapter 1.4 System of systems). Thus, precautions at individual ECPS's are necessary to enable cost efficient and trustworthy integration into SoS.
Therefore, it is essential to tackle these challenges by good engineering practices: (i) providing sets of recommended code and (system to system) interaction patterns; (ii) avoiding anti-patterns; and (iii) ensuring there is a methodology to support the integration from which the engineers of such systems can benefit. This implies aiming to resolve and pre-empt as many as possible of the integration and orchestration challenges at the platform design level. It also involves distribution of concerns to the sub systems in the SoS or IoT. Followed by automated engineering processes applying the patterns and dealing with the concerns in standardised ways. Besides this, it is necessary to facilitate communication between different stakeholders to emphasise the need for quality properties of ECPS, and to enable (automated) mechanisms that raise concerns sufficiently early to be prevented, while minimising potential losses.

On the development level, it is key to enhance the existing software systems development methodologies to support automatic engineering, also to automate the validation and verification processes for new features as they are being introduced into the system. This might need the use of AI in the validation and verification process. At this level, it is also necessary to use software system architectures in the automation of verification and other engineering practices, to manage the complexity that arises from such integration efforts (also see Major Challenge 3 below).

Artificial Intelligence and Machine Learning

Progress in AI keeps being fast-paced. While typical ECPS-needs such as Computer Vision are dominated by AI since quite some time, recent progress in Large Language Models (LLMs) opened the door to serious AI-assisted engineering tools. Hence, AI is on the way to becoming an important tool for the engineering of Embedded Software, while being an essential part of Embedded Software at the same time.

AI in Embedded Software. Simply speaking, AI lets us implement functions we don't need to understand as they are learned automatically. This has advantages and drawbacks. We take advantage of this property in tiny sensors, where AI (e.g. TinyML) automatically approximates a mapping of property changes of some material to a measurement value, in prediction, control or virtual sensing, where AI has learned to interpret time series, images, or other forms of data, and many other tasks, including detection and tracking of objects. AI functions can also serve as (automated) abstractions. With this versatility, wide-spread adoption, and a potential non-understanding of what was learned, come challenges like fitting the AI functionality to the available resources (cf. RISC-V with research on AI accelerators), testability / explainability / trustworthiness / integration in safety critical applications, real-time performance, updateability / continuous learning / maintainability, and more. In general, AI in Embedded Software faces all the same challenges standard Embedded Software has, some in a more demanding form of course. For example, using sub-symbolic AI in a safety-critical context is without a clear-cut solution regarding the trustworthiness issues. To compensate for these, extensive monitoring, and the implementation of certain architectures (e.g. Simplex) might be necessary. This shows that relying on AI in a system will have a big impact on the system design. Relying on AI will also have a massive impact on verification planning and software engineering as, e.g. proper training and validation data sets need to be provided.

AI for Embedded Software engineering. Software engineering is the discipline of building software in a proper way. AI can help with that, and AI functionality has been explored in tools for software engineering for some time. However, up to now AI-algorithms were mostly used in tools for verification (automated test case generation, checking) and less so in others. Recent developments in LLMs demonstrate the huge potential for using AI in areas like model or code generation from natural language, refactoring or "rejuvenation" of legacy software, porting software while preserving investment, or automatically adapting software to different settings/platforms. Like before, the major challenge is that there is no guarantee on the correctness or fit-for-purpose on the output of these LLMs, which is a major research opportunity at the same time. In future, AI is likely to replace many of the software-coding activities done by engineers today. Also, domain specific models will speed-up software design & generation by a considerable amount, and AI will support engineers in system understanding and in mastering complexity.
Summing up, AI hasn't yet shown its full potential for the use in and around Embedded Software – there are still many challenges left as demanding research topics. However, based on the current state-of-art and results, it is clear that AI will be a core part of future Embedded Software and ECPS.

The RISC-V instruction set architecture (ISA) is the fifth version of the Berkeley ISA that has seen exponential commercial and academic adoption in the last 10 years thanks to its open-source nature, as well as its modularity, extendibility, and simple architecture. Many commercial users adopt RISC-V in their System-On-Chips, from both open-source repositories (where the Register-Transfer-Level description of the RISC-V CPU is published), or from closed-source IP vendors. The increase in silicon devices together with the need for digital sovereignty caused many national and international organizations to take action, and RISC-V is a main actor of this revolution. For example, the European Commission[11] is building an open-source ecosystem to expand its innovation on RISC-V to compete with existing commercial alternatives, covering hardware design and system-on-chips, all the way to electronic design automation tools and the full software stack. In this scenario, OpenHW Group[12] and the Eclipse Foundation[13] play a key role in developing open-source, high-quality, silicon-proven RISC-V IPs under a permissive license. Both the not-for-profit foundations are driven by their members, who invested in RISC-V and open-source-based System-On-Chip. The modularity and extendability of the RISC-V ISA allow users to design their own architecture to meet the ultra-low-power and energy-efficient edge-computing devices constraints, as well as high-performance server machines requirements. For example, the "RVV RISC-V Vector" ISA extension allows to process multiple data concurrently, or custom extensions for security to encrypt data more efficiently. For this reason, it is crucial to have a holistic software stack to cope with the wide range of applications, taking into account the deployment of efficient applications leveraging the different ISA extensions and computer architectures.

1.3.6.1.6. Key focus areas

The key focus areas identified for this challenge include the following:

- Continuous integration of embedded software:
    - Model based design and digital twins to support system integration (HW/SW) and HW/SW co-development (increasingly new technologies have to be integrated).
    - Applying automation of engineering, taking architecture, platforms and models into account.
    - Virtualisation and simulation as tools for managing efficient integration and validation of configurations, especially for shared resources and other dependability issues.

---

[11] R. Kazman, K. Schmid, C. B. Nielsen and J. Klein, "Understanding patterns for system of systems integration," 2013 8th International Conference on System of Systems Engineering, 2013, pp. 141-146, doi: 10.1109/SYSoSE.2013.6575257

[12] https://digital-strategy.ec.europa.eu/en/library/recommendations-and-roadmap-european-sovereignty-open-source-hardware-software-and-risc-v

[13] https://www.openhwgroup.org/projects/

    - Application of integration and orchestration practices to ensure standard solutions to common integration problems.
    - Integration and orchestration platforms and separation of concerns in SoS and IoT.
    - Enabling reliable and safe continuous SW delivery to already working devices.

- Verification and validation of embedded software:
    - (Model) test automation to ensure efficient and continuous integration of CPSs.
    - Enabling secure and safe updates (cf. Major Challenge 3) and extending useful life (DevOps).
    - Continuous integration, verification and validation (with and without AI) enabling continuous certification with automated verification & validation (especially the focus on dependability), using model-based design technologies and digital twins; also when SoS and IoT are involved.
    - Certification of safety-critical software in CPSs.

1.3.6.3 Major Challenge 3: Lifecycle management

1.3.6.3.1 State of the art

Complex systems such as airplanes, vehicles and medical equipment are expected to have a long lifetime, often up to 30 years. The cost of keeping these embedded systems up to date, making them relevant for the everyday challenges of their environment is often time-consuming and costly. This is becoming more complex due to the fact that most of these systems are cyber-physical systems, meaning that they link the physical world with the digital world, and are often interconnected with each other or to the internet. With more and more functionalities being realized by embedded software, over-the-air updates – i.e. deploying new, improved versions of software-modules unto systems in the field – become an increasingly relevant topic. Apart from updates needed for error and fault corrections, performance increases and even the implementation of additional functionalities – both optional or variant functionalities that can be sold as part of end-user adaptation as well as completely new functionalities that are needed to respond to newly emerging environmental constraints (e.g. new regulations, new features of cooperating systems). Such update capabilities perfectly fit and are even required for the 'continuous development and integration' paradigm.

Embedded software also must be maintained and adapted over time, to fit new product variants or even new product generations and enable updateability of existing systems. If this is not effectively achieved, the software becomes overly complex, with prohibitively expensive maintenance and evolution, until systems powered by such software are no longer sustainable.

We must break this vicious cycle and find new ways to create software that is long-lasting and which can be cost-efficiently evolved and migrated to use new technologies. Practical challenges that require significant research in software sustainability include: (i) organisations losing control over software; (ii) difficulty in coping with modern software's continuous and unpredictable changes; (iii) dependency of software sustainability on factors that are not purely technical; (iv) enabling "write code once and run it anywhere" paradigm.

1.3.6.3.2 Vision and expected outcome

As software complexity increases, it becomes more difficult for organisations to understand which parts of their software are worth maintaining and which need to be redeveloped from scratch. Therefore, we need methods to reduce the complexity of the software that is worth maintaining and extracting domain knowledge from existing systems as part of the redevelopment effort. This also relates to our inability to monitor and predict when software quality is degrading, and to accurately estimate the costs of repairing it. Consequently, sustainability of the software is often an afterthought. This needs to be flipped around – i.e. we need to design "future-proof" software that can be changed efficiently and effectively, or at least platforms for running software need to either enable this or force such way of thinking.

As (embedded) software systems evolve towards distributed computing, SoS and microservice-based architectural paradigms, it becomes even more important to tackle the challenges of integration at the higher abstraction levels and in a systematic way. Especially when SoS or IoT is involved, it is important to be able to separate the concerns over the subsystems.

The ability of updating systems in the field in a way that safety of the updated systems as well as security of the deployment process is maintained will be instrumental for market success of future ECPS. Over-The-Air solutions become key enablers to this regard, especially for distributed systems, and they will have to cover the different paces at which HW and SW evolve, determining when updates become necessary. Edge-to-cloud continuum represents an opportunity to create software engineering approaches and engineering platforms that together enable deployment and execution of the same code anywhere on this computing continuum.

The ability of keeping track of system parameters like interface contracts and composability requires a framework to manage these parameters over the lifetime. This will enable the owner of the system to identify at any time how the system is composed and with what functionality. To this regard, the onboarding process of the constituent systems becomes a crucial phase to maintain the desired levels of security and safety: SoS integration platforms should provide solid onboarding procedures that guarantee no compromised HW and SW become part of the SoS. The adoption of block chain technologies, digital contracts and security certificates, etc. could prevent similar situations, which could impact not only the SoS during operation but also the entire supply chain associated to it. The onboarding phase should also be automated to increase security levels and ensure scalability.

Instead of focusing just on the efficiency of embedded software engineering, we already see that the field is evolving into the direction of cyber physical systems (cf. Chapter 2.3 **Architecture and Design: Methods and Tools**), and software is one element of engineering.

Many software maintenance problems are not actually technical but people problems. There are several socio-technical aspects that can help, or hinder, software change. We need to be able to organise the development teams (e.g. groups, open-source communities) in such a way that it embraces change and facilitates maintenance and evolution, not only immediately after the deployment of the software but for any moment in the software lifecycle, for the decades that follow, to ensure continuity. We need platforms that are able to run code created for different deployment infrastructures, without manual configuration.

The expected outcome is that we are able to keep embedded systems relevant and sustainable across their complete lifecycle, and to maintain, update and upgrade embedded systems in a safe and secure, yet cost-effective way.

1.3.6.3.3 Key focus areas

The key focus areas identified for this challenge include the following.

- Rejuvenation of systems:
    - Software legacy and software rejuvenation to remove technical debt (e.g. software understanding and conformance checking, automatic redesign and transformation).
    - Continuous platform-agnostic integration, deployment and migration.
    - End-of-life and evolving off-the-shelve/open source (hardware/software).

- Managing complexity over time:
    - Interplay between legacy software and new development approaches.
    - Vulnerability of connected systems.
    - Continuous certification of updates in the field (reduce throughput time).
    - Intelligent Diagnostics of systems in the field (e.g. guided root cause analysis).

- Managing configurations over time:
    - Enable tracking system configurations over time.
    - Create a framework to manage properties like composability and system orchestration.

- Evolvability of embedded software:
    - Technology, including automation of engineering and the application of integration and orchestration platforms, for keeping systems maintainable, adaptable and sustainable considering embedded constraints with respect to resources, timing and cost: new functionalities enabling and facilitating secure and automated onboarding processes, OTA software maintenance (see also the SoS chapter), ….
    - Embedded software architectures to enable SoS.

1.3.6.4 Major Challenge 4: Embedded Artificial Intelligence

1.3.6.4.1 State of the art

For various reasons – including privacy, energy efficiency, latency and the increasing necessity of smart data analytics on site – processing and artificial intelligence are moving towards the edge (edge computing), forcing the software stacks of embedded systems to coherently evolve supporting these new computing paradigms. As detailed in the Chapter "Edge Computing and Embedded Artificial Intelligence", non-functional constraints of embedded systems, such as timing, energy consumption, low memory and computing footprint, being tamperproof, etc., need to be taken into account compared to software with similar functionalities when migrating these from the cloud to the edge. Furthermore, the Quality, Reliability, Safety and Security Chapter states that key quality properties when embedding of AI components in digitalized ubiquitous systems are determinism, understanding of nominal and degraded behaviours of the system, their certification and qualification, and clear liability and responsibility chains in the case of accidents. When engineering software contains AI-based solutions, it is important to understand the challenges that such solutions introduce. Indeed, AI contributes to address challenges of embedded software, but it does not define them exclusively itself, as quality properties of embedded software depend on integration of AI-based components with other software components.
For efficiency reasons, very intensive computing tasks (such as those based on deep neural networks, DNNs) are being carried out by various accelerators embedded in systems on a chip (SoCs). Although the "learning" phase of a DNN is still mainly done on big servers using graphics processing units (GPUs), local adaptation is moving to edge devices. Also, LLMs are experiencing a similar evolution towards the edge. Alternative approaches, such as federated learning, allow for several edge devices to collaborate in a more global learning task. Therefore, the need for computing and storage is ever-increasing, and is reliant on efficient software support.

The "inference" phase (i.e. the use after learning) is also requiring more and more resources because neural networks are growing in complexity exponentially. Once carried out in embedded GPUs, this phase is now increasingly performed on dedicated accelerators. Most middle and high-end smartphones have SoCs embedding one of several AI accelerators, as well as Mx Apple processors family for laptops, tablets and future wearables – for example, the Nvidia Jetson Xavier NX is composed of six Arm central processing units (CPUs), two inference accelerators, 48 tensor cores and 384 Cuda cores. **Obtaining the best of the heterogeneous hardware is a challenge for the software, and the developers should not have to be concerned about where the various parts of their application are running.**

Once developed (on servers), a neural network has to be tuned for its embedded target by pruning the network topology using less precision for operations (from floating point down to 1-bit coding) while preserving accuracy. This was not a concern for the "big" AI development environment providers (e.g. Tensorflow, PyTorch, Caffe2, Cognitive Toolkit) until recently. This has led to the development of environments[14] designed to optimise neural networks for embedded architectures[15] to move towards the Edge.

Most of the time the learning is done on the cloud. For some applications/domains, making a live update of the DNN or LLM characteristics is a sought-after feature, including all the risks of security, interception. Imagine the consequences of tampering with the DNN or LLM used for a self-driving car! A side-effect of DNN or LLM is that intellectual property is not in a code or algorithm, but rather lies in the network topology and its weights, and therefore needs to be protected.

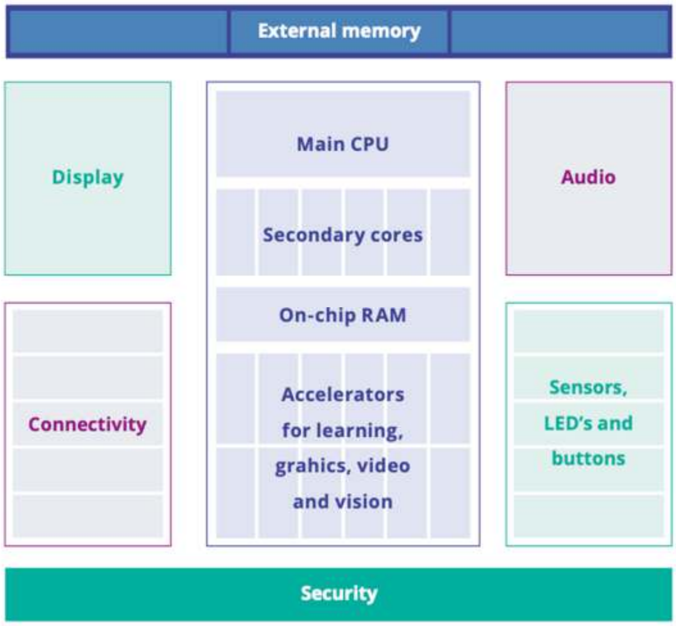


Figure 1.3.5 - Data analytics and Artificial Intelligence require dedicated embedded hardware architectures

1.3.6.4.2 Vision and expected outcome

European semiconductor companies lead a consolidated market of microcontrollers and low-end microprocessors for embedded systems, but are increasing the performance of their hardware, mainly driven by the automotive market and the increasing demand for more performing AI for advanced driver-assistance systems (ADAS) and self-driving vehicles. With edge computing and embedded AI this trend has extended to other vertical domains that are key for Europe. European semiconductor companies are also moving towards greater heterogeneity by adding specialised accelerators. On top of this, Quality, Reliability, Safety and Security Chapter lists personalization of mass products and resilience to cyber-attacks, as the key advantage and the challenge characterizing future products. Embedded software

[14] E.g. the Eclipse Aidge platform aims to facilitate the design and optimization of neural networks for easy deployment of embedded AI applications on different hardware targets with low power consumption and small memory footprint.

[15] https://www.eclipse.org/

needs to consider these and find methods and tools to manage their effects on quality properties of software that integrates them. Also, embedded software engineering will need to ensure interoperability between AI-based solutions and non-AI parts.

In this context, there is a need to provide a programming environment and libraries for the software developers. A good example here is the interchange format ONNX, an encryption format for protection against tampering or reverse engineering that could become the foundation of a European standard. Beside this, we also need efficient libraries for signal/image processing for feeding data and learning into the neural network, abstracting from the different hardware architectures. These solutions are required to be integrated and embedded in ECPS, along with significant effort into research and innovation in embedded software.

1.3.6.4.3 Key focus areas

The key focus areas identified for this challenge include the following.

- Federated and distributed learning:
  - Create federated learning at the edge in heterogeneous distributed systems (analysis, modelling and information gathering based on local available information).
  - Federated intelligence at the edge (provide context information and dependability based on federated knowledge).

- Embedded Intelligence:
  - Create a software AI framework to enable reflecting and acting on the systems own state.
  - Dynamic adaption of systems when environment parameters and sensors like IoT devices are changing.

- Data streaming in constraint environments:
  - Feed streaming data into low-latency analysis and knowledge generation (using context data to generate relevant context information).

- Embedding AI accelerators:
  - Accelerators and hardware/software co-design to speed up analysis and learning (e.g. patter analysis, detection of moves (2D and 3D) and trends, lighting conditions, shadows).
  - Actual usage-based learning applied to accelerators and hardware/software co-design (automatic adaptation of parameters, adaptation of dispatch strategies, or use for new accelerators for future system upgrades).

1.3.6.5 Major Challenge 5: Support for sustainability by embedded software

1.3.6.5.1 State of the art

The complete power demand in the whole ICT market currently accounts from 5% to 9% of the global power consumption[16]. The ICT electricity demand is rapidly increasing and it could go up to nearly 20% in 2030. Compared to estimated power consumption of future large data centres, embedded devices may seem to be a minor problem. However, when the devices are powered by batteries, they still have a significant environmental impact. Energy efficient embedded devices produce less hazardous waste and last longer without the need for replacement.

The growing demand for ultra-low power electronic systems has motivated research into device technology and hardware design techniques. Experimental studies have proven that the hardware innovations for power reduction can be fully exploited only with proper design of the upper layer software. Partitioning hardware enables smart power-up/power-down strategies. Together with support through resource-aware algorithms, this could lead to significant energy savings. The same applies to software power and energy modelling and analysis: the first step towards the energy reduction is complex due to the inter- and intra-dependencies of processors, operating systems, application software, programming languages and compilers. Software design and implementation should be viewed from a system energy conservation angle rather than as an isolated process.

For sustainability, it is critical to understand quality properties of software. These include in the first place power consumption, and then other related properties (performance, safety, security, and engineering-related effort) that we can observe in the context of outdated or inadequate software solutions and indicators of defected hardware. Power reduction strategies are mainly focusing on processing, storage, communication, and sometimes on other (less intelligent) equipment.

For the future embedded software developers, it is crucial to keep in touch with software development methodologies focused on sustainability, such as green computing movement, resource-aware computing and, more generally, sustainable programming techniques. In the domain of embedded software, examples include the estimation of the remaining useful life of the device, the network traffic and latency time optimization, the process scheduling optimization or energy efficient workload distribution, the management of HW and SW resources oriented to energy saving, the correct use of software abstraction, etc.

1.3.6.5.2 Vision and expected outcome

The concept of sustainability is based on three main pillars: ecological, economic and social. The ideal environmentally sustainable (or green) software in general requires as little hardware as possible, it is efficient in power consumption, and its usage leads to minimal waste production. Embedded software designed to be adaptable for future requirements

---

[16] Such as N2D2, https://github.com/CEA-LIST/N2D2

without the need to be replaced by a completely new product is an example of environmentally, economically, and socially sustainable software.

To reach the sustainability goal, the embedded software design shall focus also on energy-efficient design methodologies and tools, energy efficient and sustainable techniques for embedded software and systems production and to the development of energy- and resource-aware applications and frameworks for embedded systems, edge computing, embedded intelligence and their applications.

It is evident that energy/power management has to be analysed with reference to the context, to the underlying hardware resources and the overall system functionalities. The coordinated and concentrated efforts of a system architect, hardware architect and software architect should help to introduce energy-efficient systems (cf. Chapter 2.3, Architecture and Design: Methods and Tools). The tight interplay between energy-oriented hardware, energy-aware and resource-aware software calls for innovative structural, functional and mathematical models for analysis, design and run-time. Model-based software engineering practices, supported by appropriate tools, will definitely accelerate the development of modern complex systems operating under severe energy constraints. It is crucial to notice the relationship between power management and other quality properties of software systems (e.g. under certain circumstances it is adequate to reduce the functionality of software systems by disabling certain features, which results in significant power savings). From a complementary perspective, when software is aware of the available hardware resources and their energy profile, it enables power consumption optimisation and energy saving, being able to configure the hardware resources, to activate/deactivate specific hardware components, increase/decrease the CPU frequency according to the processing requirements, partition, schedule and distribute tasks.

Therefore, in order to enable and support sustainability through software, software solutions need to be reconfigurable in the means of their quality. There must exist strategies for HW/SW co-design and accelerators to enable such configurations, and the entire integrated development environment should facilitate a rational use of abstraction (abstraction simplifies software development but increases the energy consumption) and supports HW and SW resources usage optimisation from the energy perspective directly from the programming language and through the compiler, linker, assembler, etc. For this to be possible, software systems need to be accompanied with models of their quality properties and their behaviour, including the relationship between power consumption and other high level quality properties. This will also enable balancing mechanisms between local and remote computations to reduce communication and processing energy consumption.

Models (digital twins) should be aware of energy use, energy sources, HW and SW resources and their sustainability profile. An example of this in SoS are solar cells that give different amounts of energy dependent on time of day and weather conditions.

1.3.6.5.3 Key focus areas

The following key focus areas have been identified for this challenge:

- Resource-aware software engineering.
- Tools and techniques enabling the energy-efficient and sustainable embedded software design.
- Development of energy-aware and sustainable frameworks and libraries for embedded software key application areas (e.g. IoT, Smart Industry, wearables, etc.).

- Management of computation power on embedded hardware:
  - Management of energy awareness of embedded hardware, embedded software with respect to, amongst others, embedded high-performance computing (HPC).
- Composable efficient abstractions that drive sustainable solutions while optimising performance:
  - Enabling resource-aware computing.
  - Enabling technologies for the second life of (legacy) cyber-physical systems.
  - Establish relationships between power consumption and other quality properties of software systems, including engineering effort (especially in cases of computing-demanding simulations).
  - Digital twins can support the management of quality properties of software with the goal of reducing power consumption, as the major contributing factor to the green deal, enabling sustainability.

1.3.6.6 Major Challenge 6: Software reliability and trust

1.3.6.6.1 State of the art

Two emerging challenges for reliability and trust in ECPS relate to computing architectures and the dynamic environment in which ECPS exist. The first challenge is closely related to the end of Dennard scaling[17]. In the current computing era, concurrent execution of software tasks is the main driving force behind the performance of processors, leading to the rise of multicore and manycore computing architectures. As the number of transistors on a chip continues to increase (Moore's law is still alive), industry has turned to a heavier coupling of software with adequate computing hardware, leading to heterogeneous architectures. The reasons for this coupling are the effects of dark silicon[18] and better performance-to-power ratio of heterogeneous hardware with computing units specialised for specific tasks. The main challenges for using concurrent computing systems in embedded systems remain: (i) hard-to-predict, worst-case execution time; and (ii) testing of concurrent software against concurrency bugs[19].

The second challenge relates to the dynamic environment in which ECPS execute. On the level of systems and SoS, architectural trends point towards platform-based designs – i.e. applications that are built on top of existing (integration and/or middleware) platforms. Providing a standardised "programming interface" but supporting a number of constituent

---

[17] https://www.enerdata.net/publications/executive-briefing/between-10-and-20-electricity-consumption-ict-sector-2030.html

[18] John L. Hennessy and David A. Patterson. 2019. A new golden age for computer architecture. Commun. ACM 62, 2 (February 2019), 48–60. DOI: https://doi.org/10.1145/3282307

[19] Hadi Esmaeilzadeh, Emily Blem, Renee St. Amant, Karthikeyan Sankaralingam, and Doug Burger. 2011. Dark silicon and the end of multicore scaling. In Proceedings of the 38th annual international symposium on Computer architecture (ISCA '11). Association for Computing Machinery, New York, NY, USA, 365–376. DOI: https://doi.org/10.1145/2000064.2000108

subsystems that is not necessarily known at design time, and embedding reliability and trust into such designs, is a challenge that can be solved only for very specialised cases. The fact that such platforms – at least on a SoS level – are often distributed further increases this challenge.

On the level of systems composed from embedded devices, the most important topics are the safety, security, and privacy of sensitive data. Security challenges involve: (i) security of communication protocols between embedded nodes, and the security aspects on the lower abstraction layers; (ii) security vulnerabilities introduced by a compiler[20] or reliance on third-party software modules; and (iii) hardware-related security issues[21]. It is necessary to observe security, privacy and reliability as quality properties of systems, and to resolve these issues on a higher abstraction level by design[22], supported by appropriate engineering processes including verification (see Chapters 2.3 and 2.4).

1.3.6.6.2 Vision and expected outcome

European industry today relies on developed frameworks that facilitate production of highly complex embedded systems (for example, AUTOSAR in the automotive industry).

The ambition here is to reach a point where such software system platforms are mature and available to a wider audience. These platforms need to enable faster harvesting of hardware computing architectures that already exist and provide abstractions enabling innovators and start-ups to build new products quickly on top of them. For established businesses, these platforms need to enable shorter development cycles while ensuring their reliability and providing means for verification & validation of complex systems. The purpose of building on top of these platforms is ensuring, by default, a certain degree of trust for resulting products. This especially relates to new concurrent computing platforms, which hold promise of great performance with optimised power consumption. Recent developments in programming languages - such as Rust - look promising, as they aim to solve some of these inherited problems by default, based on available programming language constructs.

Besides frameworks and platforms that enable easy and quick development of future products, the key enabler of embedded software systems is their interoperability and openness. In this regard, the goal is to develop and make software libraries, software frameworks and reference architectures which need to ensure, by design, the potential for monitoring, verifying, testing and auto-recovering of embedded systems. That enables

---

[20]F. A. Bianchi, A. Margara and M. Pezzè, "A Survey of Recent Trends in Testing Concurrent Software Systems," in IEEE Transactions on Software Engineering, vol. 44, no. 8, pp. 747-783, 1 Aug. 2018, doi: 10.1109/TSE.2017.2707089.

[21]V. D'Silva, M. Payer and D. Song, "The Correctness-Security Gap in Compiler Optimization," 2015 IEEE Security and Privacy Workshops, 2015, pp. 73-87, doi: 10.1109/SPW.2015.33.

[22]Moritz Lipp, Vedad Hadžić, Michael Schwarz, Arthur Perais, Clémentine Maurice, and Daniel Gruss. 2020. Take A Way: Exploring the Security Implications of AMD's Cache Way Predictors. In Proceedings of the 15th ACM Asia Conference on Computer and Communications Security (ASIA CCS '20). Association for Computing Machinery, New York, NY, USA, 813–825. DOI:https://doi.org/10.1145/3320269.3384746

interoperability and integration of products developed on distributed computing architectures available to a wider audience. One of the emerging trends to help achieving this is the use of digital twins. Digital twins are particularly suitable for the verification of safety-critical software systems that operate in dynamic environments. However, development of digital twins remains an expensive and complex process, which has to be improved and integrated as part of the standard engineering processes (see Major Challenge 2 in Chapter 2.3).

We envision an open marketplace for software frameworks, middleware, and digital twins that represents a backbone for the future development of products. While such artefacts need to exploit the existing software stacks and hardware, they also need to support correct and high-quality software by design. Special attention is required for Digital Twin simulations of IoT devices to ensure reliability and trust in operating in real life.

1.3.6.6.3 Key focus areas

Focus areas of this challenge are related to quality aspects of software. For targets such as new computing architectures and platforms, it is crucial to provide methodologies for development and testing, as well as for the team development of such software. These methodologies need to take into account the properties, potentials and limitations of such target systems, and support developers in designing, analysing and testing their implementations. As it is fair to expect that not all parts of software will be available for testing at the same time, it is necessary to replace some of the concurrently executing models using simulation technologies. Finally, these achievements need to be provided as commonly available software modules that facilitate the development and testing of concurrent software.

New concepts for programming languages, such as Rust, by their default design resolve some of the listed issues in developing computing architectures. For example, one of the main goals of Rust is handling concurrent and parallel programming in a safe and efficient way[23]. New concepts in the area of programming languages need to balance between several factors. From one side, new programming languages need to offer higher productivity to engineers. Programming languages need to enable more efficient collaboration between engineers by being more suitable to higher level architectural thinking that prioritise decoupled development of individual software components. Furthermore, traditional programming languages, such as C and C++, are challenging for static analysis due to undecidability of their statements[24] and in general have huge test space making them susceptible to security issues. Finally, new languages need to solve by design complex problems in programming, such as concurrency and parallelism. On the other hand, it is necessary to minimize overhead of

[23] https://doc.rust-lang.org/book/ch16-00-concurrency.html
[24] Michael Hind. 2001. Pointer analysis: haven't we solved this problem yet? In Proceedings of the 2001 ACM SIGPLAN-SIGSOFT workshop on Program analysis for software tools and engineering (PASTE '01). Association for Computing Machinery, New York, NY, USA, 54–61. https://doi.org/10.1145/379605.379665

abstractions in new programming languages (performance[25] and energy[26] [27] overhead). Besides creating such languages or extending the existing ones with news features, it is still necessary to provide methodologies that will guide industry to migrate their code bases to new programming constructs, or at least ensure co-existence and interaction between new and old code bases. With the introduction of AI into software engineering, we are looking for programming languages that will facilitate a new AI-assisted[28] software development approach. That means programming languages that engineers can use to easily and as deterministically as possible express their intentions, while those languages are suitable for AI to enable code generation and automatization of validation, verification, and testing activities. Furthermore, we hope to have in the future such programming languages that facilitate, with the assistance from AI, analysis of quality properties of whole software stacks (e.g., WCET analysis, safety analysis[29]).

The next focus area is testing of systems against unexpected uses, which mainly occurs in systems with a dynamic execution environment. It is important here to focus on testing of self-adapting systems where one of the predominant tools is the simulation approach, and more recently the use of digital twins.

However, all these techniques are not very helpful if the systems are not secure and reliable by design. Therefore, it is necessary to investigate platforms towards reliability, security and privacy, with the following challenges.

- Reliable software on new hardware including edge, fog and cloud processing: (co) verification of distributed, also heterogenous systems.
- Verification and validation of ML models.
- Robustness against unexpected uses:
    - Trustworthy, secure, safe, privacy-aware.
    - Validating self-adapting systems for example through simulation.
- Security and privacy as a service:
    - To become part of the software architecture.
    - Means and techniques for continuous system monitoring and self-monitoring.

---

[25] There’s plenty of room at the Top: What will drive computer performance after Moore’s law? E. Leiserson et all, Science 05 Jun 2020: Vol. 368, Issue 6495, DOI: 10.1126/science.aam9744

[26] Rui Pereira, Marco Couto, Francisco Ribeiro, Rui Rua, Jácome Cunha, João Paulo Fernandes, and João Saraiva. 2017. Energy efficiency across programming languages: how do energy, time, and memory relate? In Proceedings of the 10th ACM SIGPLAN International Conference on Software Language Engineering (SLE 2017). Association for Computing Machinery, New York, NY, USA, 256–267. https://doi.org/10.1145/3136014.3136031

[27] Rui Pereira, Marco Couto, Francisco Ribeiro, Rui Rua, Jácome Cunha, João Paulo Fernandes, João Saraiva, Ranking programming languages by energy efficiency, Science of Computer Programming, Volume 205, 2021

[28] Russo, Daniel & Baltes, Sebastian & van Berkel, Niels & Avgeriou, Paris & Calefato, Fabio & Cabrero-Daniel, Beatriz & Catolino, Gemma & Cito, Jürgen & Ernst, Neil & Fritz, Thomas & Hata, Hideaki & Holmes, Reid & Izadi, Maliheh & Khomh, Foutse & Kjærgaard, Mikkel & Liebel, Grischa & Lafuente, Alberto & Lambiase, Stefano & Maalej, Walid & Vasilescu, Bogdan. (2024). Generative AI in Software Engineering Must Be Human-Centered: The Copenhagen Manifesto. Journal of Systems and Software. 216. 112115. 10.1016/j.jss.2024.112115.

[29] https://hightec-rt.com/rust

1.3.6.7 Major Challenge 7: Hardware virtualization for efficient SW engineering

1.3.6.7.1 State of the art

Hardware virtualisation provides efficient abstraction from the physical hardware, thus allowing to decouple software engineering lifecycles from the underlying hardware. This is usually implemented either by full-scale hypervisors like VMware and KVM, or via containerisation as combination with the hosting operating system, e.g. Docker and Kubernetes. It enhances efficiency through resource isolation and allocation, and eases deployment by offering lightweight, portable, and scalable software packaging. Additionally, hardware-assisted virtualisation extensions such as Intel VT-x and AMD-V have improved performance and security by offloading virtualisation tasks to the hardware.

Hardware virtualisation is expected to offer numerous benefits, including increased flexibility, portability, and security for safety-critical and non-safety-critical applications.

1.3.6.7.2 Vision and expected outcome

- **Standardized Abstraction Models** for hardware components to enable cross-platform compatibility, fostering ease of development and integration.
- **Timing Models**: Developing reliable timing models to predict and verify real-time behaviour accurately.
- Unified **APIs** for different hardware components to promote interoperability.
- **Performant run-time environments** supporting shared memory access control, leveraging hypervisor technologies like XEN and KVM, as well as Bytecode/WebAssembly technologies.
- Virtualisation of communication (**Virtual Networks**) supporting Quality of Service (QoS) mechanisms to guarantee real-time communication in mixed-criticality environments.
- Hardware Abstraction for **Sensors and Actuators**, to achieve their interchangeability.

1.3.6.7.3 Key focus areas

To realize the potential of hardware virtualisation, research and development will focus on three aspects.

- Standard development methods and frameworks for the development of hardware abstractions, integrated with existing tools.
- Verification and validation frameworks, supported by automation, which allow for the validation of applications within virtualisation, as well as the validation of specific target systems, to confirm performance and timeliness.
- Run-time environments for safety-critical applications.

To cover these aspects, it is necessary to understand what they aim to achieve in software system engineering terms. To facilitate software development in virtual environments, the most common requirement in industry is to enable functional testing on virtual hardware. Such testing must provide fast execution feedback suitable for continuous engineering. Currently, there already exist techniques that try to achieve this[30]. What we also need is higher integration of such virtual platforms in common engineering development methodologies (e.g., Agile). To achieve other goals from the list above (e.g., performance and temporal properties evaluation), it is necessary to have more complex, detailed models of virtual hardware. Examples of such models currently exist[31], but are extremely slow to execute. Besides integration of virtual platforms into existing methodologies and making execution of these models faster, it is necessary to observe the challenge around this effort in a wider scope. Development of virtual prototypes is not easy and often requires significant effort and investment. Moreover, management of platforms on which these prototypes execute is also a complex task (e.g. using BlueChi[32]). Finally, management of models, their variations, and integration activities with software repositories and test suites is challenging and time consuming. Therefore, we need approaches that facilitate faster development of virtual platforms, infrastructure around using them in software engineering, and management of both the models and the infrastructure. With these efforts, we hope to migrate a larger part of the development of embedded software systems to the cloud, facilitating collaboration and enabling higher engineering efficiency.

[30] https://newsroom.arm.com/blog/isa-parity
[31] https://www.gem5.org/
[32] https://projects.eclipse.org/projects/automotive.bluechi

1.3.5. TIMELINE

The following table illustrates the roadmaps for Embedded Software and Beyond. The assumption is that on the topics listed, that technology should be ready (TRL 8–9) in the respective time-frames.

MAJOR CHALLENGE	TOPIC	SHORT TERM (2025-2029)	MEDIUM TERM (2030-2034)	LONG TERM (2035 and beyond)
Major Challenge 1: Efficient engineering of embedded software	Topic 1.1: Modelling-based software engineering	Model-based software engineering enabling systems to become part of SoS	Model inference to enable re-use of existing subsystems in SoS	Model-based testing taking re-use of uncontrolled SoS into account
	Topic 1.2 Digital twinning	Virtualisation of legacy systems	support virtual integration testing across variants	Support/allow (re)certification of systems via digital twins
	Topic 1.3: Constraint environments	Resource planning and scheduling Design for software evolution over time	Embedded software architectures to enable SoS	Exploiting hybrid computer platforms, including efficient software portability
	Topic 1.4: Software technology	Virtualisation as tool for efficient engineering Technology for safe and dependable software ecosystems	Interface management enabling systems to become part of SoS	Develop new software architectures for edge computing Artificial intelligence to assist and support efforts in software engineering
	Topic 1.5: Software engineering tools	Co-simulation platforms	Middleware controlling dynamically embedded (mobile) hardware solutions Compilers and link to new hardware	Programming languages for developing large-scale applications for embedded SoS

Major Challenge 2: Continuous integration and deployment	Topic 2.1: Continuous integration	DevOps modelling Virtualisation	Simulation on a virtual platform	Digital twin Model-based engineering based on digital twins
	Topic 2.2: Verification and validation	Virtualisation of test platform	Model-based testing	Integration & orchestration platforms for IoT and SoS
Major Challenge 3: Life-cycle management of embedded software	Topic 3.1: Rejuvenation of existing systems	Software legacy and software rejuvenation Design for rejuvenating systems in a later phase	End-of-life and evolving off-the-shelve/open-source solutions	The cloud-for-edge continuum - “Write once, run anywhere” on this computing continuum Composability, properties contracts and orchestration systems Interoperability: must be ensured in integration platforms
	Topic 3.2: Managing complexity over time	Diagnostics of systems in the field	Continuous certification	Interplay between legacy
	Topic 3.3: Managing Configurations over time	Full life-cycle configuration tacking	Methods and tools managing composability and system orchestration	Individualized systems configuration management
	Topic 3.4: Evolvability of embedded software	Adaptable embedded software	Dynamical embedded software	Autonomous embedded software Autonomous processes (IoT & edge embedded HW/SW co-design)
Major Challenge 4: Embedding data analytics and AI	Topic 4.1: Federated learning	Create federated learning at the edge in heterogeneous distributed systems	Federated intelligence at the edge	Safe, trustworthy & explainable AI AI is playing several key roles in innovation, e.g. as a tool for SW development/ engineering Embedded intelligence

	Topic 4.2 Embedded Intelligence	Self-reflection: software AI framework supports acting on own system state	OS support for new HW (GPU, ASIC, neuromorphic computing,…) and platforms (Edge-AI,..)	Support dynamic adaption of systems
	Topic 4.3: Data streaming in constraint environments	Feed streaming data into low-latency analysis and knowledge generation	Support processing by new HW (GPU, ASIC,…)	
	Topic 4.4: Embedding AI accelerators	Accelerators and hardware/software co-design to speed up analysis and learning	Actual usage-based learning applied for accelerators and hardware/software co-design	Use of AI in autonomous systems
Major Challenge 5: Support for sustainability by embedded software	Topic 5.1: resource-aware software engineering	Integration of green-aware aspects in software integration	Adaptive processing based on energy-awareness	
	Topic 5.2: Tools for energy efficient SW design	Rejuvenation technologies	Design for extending lifetime	Digital twins that support green deal and enable sustainability (e.g. contain power models)
	Topic 5.3: Energy aware frameworks & libraries	Support monitoring/ reporting energy production / energy profiles	Support scale-able processing depending on available energy	Energy-optimal distributed computation
	Topic 5.4: Management of computation power on embedded HW	SW/HW support for energy awareness of embedded systems	Support for embedded HPC	

	Topic 5.5: Composable efficient abstraction	Enabling technologies for the second life of (legacy) cyber-physical systems	Establish relationships between power consumption and other quality properties	
Major Challenge 6: Software reliability and trust	Topic 6.1: Reliability of software and new hardware	Code coverage of reliability tooling and porting Simulation- and mock-up- based approaches for handling concurrency	Embed reliability on software architecture level	Use of quantum computing IoT digital twin simulation Validation and verification through simulation- and mock-up-based approaches for handling concurrency
	Topic 6.2: Robustness (trustworthy, secure, safe, privacy-aware)	Trustworthy, secure, safe, privacy-aware Testing self-adapting systems using simulation	Define a maturity model for robustness of embedded software and beyond	
	Topic 6.3: Security and privacy as a service	Design for security and privacy as a service	Architecture for security and privacy as a service	
Major Challenge 7: Hardware virtualization for efficient SW engineering	Topic 7.1: development methods and frameworks for hardware abstractions	Modular building blocks available for creating abstraction layers for common multicore-CPU/SoC platforms	Design automation for abstraction layers built from formal HW description	Methodology and tools automating abstraction layer design, providing certain guarantees (safety, security, determinism etc.)
	Topic 7.2: validation of application frameworks	Support virtualized V&V of applications on abstraction layers for wide range of target systems and variants thereof	Automated validation of properties like safety, security and runtime determinism	

	Topic 7.3: Run-time environments for safety-critical applications	Highly performant & analyse-able run-time environments supporting shared memory access control	Hardware abstraction frameworks fit to be certified for safety-critical domains (automotive, aeronautics,…)	Hardware abstraction frameworks fit for certification in **highly** safety-critical applications (e.g. ASIL-D in automotive)

1.4

Foundational Technology Layers

SYSTEM OF SYSTEMS

1.4 System of Systems

1.4.1. SCOPE

The System of Systems (SoS) technology layer represents the upper layer of ECS technology stack for digitalisation solutions. This technology layer emerges from the composition of embedded and cyber-physical systems (CPS), connectivity and distributed software platforms.

In the ECS domain, a constituent system of a System of Systems (SoS) is defined as a set of embedded hardware hosting software designed to perform a particular task or solve a specific problem. A constituent system can be distributed, but from a logical/conceptual perspective it is “contained” in one unit and it is autonomous and/or independent from the other constituent systems, (i.e. it shows managerial and operational independence from any other constituent system). The complexity of these constituent systems is rapidly increasing with the development of the underlying HW/SW technologies, as well as the rising demand by the users of these systems for functional and extra-functional requirements.

According to the definition developed by Mayer, 1998[1], SoS must satisfy five characteristics: (i) the operational independence of constituent systems; (ii) the managerial independence of constituent systems; (iii) geographical distribution; (iv) emergent behaviour; and (v) evolutionary development processes. A system that does not satisfy these characteristics is not considered an SoS.

For existing systems this independence results in composing or integrating systems that were not designed together, to perform a combined task besides their ‘normal’ task. SoS engineering aims at methods and architectures to resolve this, typically addressing resource sharing, and access to data and services. Model-based techniques for the design of an SoS can be used in a similar way as for regular systems; however, the integrating systems rely on different models and paradigms. Further methodology is needed to address that systematically.

Newly developed systems must be designed such that they are prepared for forms of SoS integration. Here, model-based techniques are useful, for example, in the application of AI techniques e.g. for learning dynamically how systems must work together while increasing the semantic level of interoperability. Research should address the development of methodology and standard patterns, interfaces and artifacts for SoS that complement current methodology for system design. Focus should be on the aspects that are specific for SoS such as the mentioned independence, and the integration into an SoS: discovery and use of services, the sharing of data and resources, the support for extra-functional properties and the very late binding, to be negotiated at interfaces. Such negotiation

---

[1] Architecting Principles for Systems-of-Systems, Mark W. Maier, Systems Engineering journal, John Wiley &Sons 1998

requires predictive models that support taking sharing decisions and build on interoperability and trustworthiness.

In modern hyper-connected digital solutions, systems rarely operate independently. On the contrary, the primary added value of these digital solutions is the cooperation between heterogeneous systems to solve more complex problems by exploiting the set of multi-technology, multi-brand and even multi-domain functionalities generated by the cooperation. While talking or reading, SoS is typically pronounced entirely “System of Systems”. An SoS emerges from the composition/integration of multiple systems to perform a task or reach an objective that none of the constituent systems can perform or reach on their own. In the SoS, each constituent system is considered a “black box”: it remains operational and managerial autonomous and/or independent, relying on its own hardware, software, and networking resources, and remaining focused on its own goals. At the SoS level, the SoS evolves with components, functions and purposes added, removed, and modified, leading to an increasing dynamicity and variability along their life cycle (a life cycle that is intertwined with the life cycles of constituent systems and potentially never finishes!). The SoS structure evolves with the addition or removal of the constituent systems, which always cooperate, coordinate, and adapt to achieve the SoS goals, providing additional features to the SoS as a whole and capabilities and functionalities unavailable in the constituent systems. Having an up-to-date inventory and real time monitoring of the SoS is challenging.

Like a nervous system – i.e., partially centralized, distributed and peripheral – a software integration infrastructure is a key element of an SoS. The nervous system has an architecture, and so does an SoS. The most common architecture approach in SoS is based on SOA (Service Oriented Architecture) and micro-service from edge to cloud. Splitting such architecture into three parts; Infrastructure, Integration platform and Solution implementation, provides a logical base for what can be shared and what will be company specific, among involved stakeholders, as shown in Figure 1.4.1.

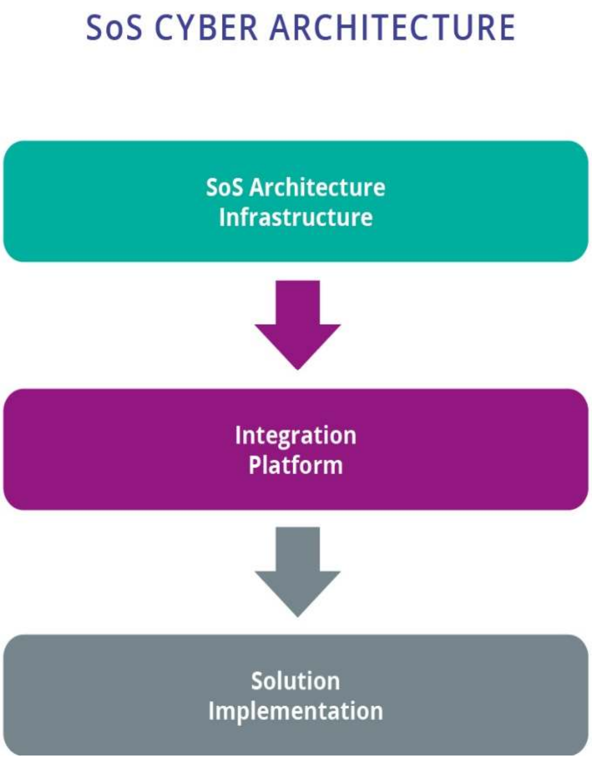


Figure 1.4.1. An SoS cyber architecture provides, based on SOA architecture, an infrastructure supporting fundamental service-oriented properties like, Look-up, Late binding and Loose coupling plus a number of support functionalities to build working solutions.

The SOA architecture infrastructure will provide e.g.:

- Fundamental SOA property support: Look-up, Late binding and Loose coupling
- Security support
- Interoperability support
- System of Systems integration support
- Basic engineering and operations support
- Model-based engineering support

Using the SoS infrastructure, SoS platforms are created by vendors and larger companies. SoS platforms will be used for solution-specific implementations, engineered, deployed and operated.

To create added value, an SoS needs to be trustworthy, and here e.g. end-to-end security issues have to be properly taken into account. A secure SoS should be able to defend against both deliberate attacks and accidental threats, and also its misuse. Moreover, it is not enough to ensure that each of the constituent systems is secure in the pre-deployment phase, but also that the evolved/composed/integrated SoS, whose exact composition may be not known in advance, is secure. Dynamic adaption to e.g. security or safety requirements and risks analysis should be considered over time in relation to emergent functionalities, properties and behaviors arising from the complex interactions among the constituents of the SoS. New methodology and tools for risk and vulnerability assessment and threat modeling are needed. Artificial intelligence, machine learning and ontology/semantic-based approaches can complement each other for improved knowledge and decision-making processes in an SoS structure. AI/ML can make predictions based on experience or training, while ontologies/semantics provide information based on reasoning which also can optimise and accelerate machine learning processes.

It is unrealistic to imagine that a single SoS infrastructure could drive an entire market because, considering the interdisciplinarity and complexity required to develop them, very seldom will a single vendor be able to provide a complete end-to-end and domain-independent solution. However, platform "competition" will at least have to identify a set of European solutions that covers key vertical domains. For key European vertical domains an SoS has to address a multitude of cross sectorial requirements like e.g. security, safety, evolution, maintenance, trustworthiness. For example, security and safety certification issues both at component, system and SoS level should be properly addressed aiming at really mitigating risks/threats in competitive scenarios, while also considering the EU Cybersecurity Certification framework.

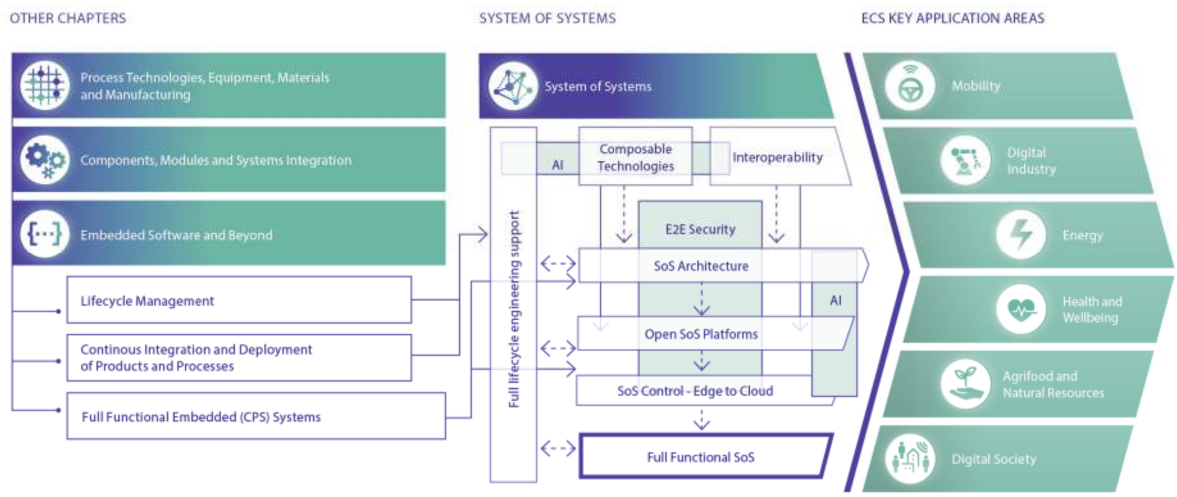


Figure 1.4.2 - Structure: System of Systems

1.4.2. Application breakthroughs

Improvements in SoS technology will have an impact on all ECS application areas. They will enable or support faster translation of ideas into economically viable solutions and might open new markets by further upscaling.

Examples of health and well-being application breakthroughs supported by SoS are:

- Interoperability of health data.
- Strengthening where and how healthcare is delivered, supporting home-based care.
- Supporting the clinical workforce and healthcare consumers to embrace technology-enabled care.

- High level of digital trust.
- Data security technology for interoperability between security hardware and software components.
- Improved integration and analysis of multimodal data.
- Integration platforms for embedded ultrasound, low-power edge computing, and AI and digital health.

For the mobility application area, the provision of EU capabilities within SoS will support breakthroughs regarding:

- Achieving the Green Deal for mobility with the 2 Zero goals of –37.5% CO_2 by 2030.
- Increased road safety through the CCAM[2] programme.
- Improve the competitiveness of the European industrial mobility digitalisation value chain.
- Ensuring inclusive mobility for persons and goods by providing mobility access to everyone, with a focus on special needs.

In the energy application domain, the provision of improved SoS capabilities and engineering efficiency will support breakthroughs regarding:

- Management of multivalent sector coupling (electricity, heating / cooling, mobility) for the future all-electric society.
- Supporting grid stabilisation by intermediate storage share of renewable energies, peak control or viability management for the increase of energy flexibility.
- Energy supply infrastructure for e-mobility, digital live, and industry 4.0.
- “Plug and play integration” of ECS into self-organised grids and multi-modal systems, real-time digital twin capability in component and complete system design (to simulate system behaviour).
- Significant reduction and recovery of losses (application and SotA-related).
- Increased functionality, reliability, and lifetime (incl. sensors & actuators, ECS HW/SW, semiconductor power devices, artificial intelligence, machine learning, monitoring systems, etc.).
- Safety and security issues of self-organised grids and multi-modal systems through smart edge devices and high-level IT security (resilient communications and trustworthy AI).
- Optimisation of applications and exploitation of achieved technology advances in all areas where electrical energy is consumed.
- Energy technologies in the circular economy approach: predictive and condition-based maintenance with repair and recycle capabilities.
- Aligning with standardisation of different energy systems.

In the industry and agrifood application domains, the provision of advanced SoS architectures, platforms and engineering automation will support the EU regarding:

- Intelligent control room systems to enable correlations between machine malfunctions and load parameters to be detected immediately, thereby enabling

[2] https://www.ccam.eu

maintenance work to be carried out early and on schedule, with a reduction in costly downtimes.

- Food industry imposes specific requirements (e.g. in food processing) that may take advantage of smart (bio-)sensing for high-quality monitoring to reduce the amount of water and chemicals used in such processes, and to prevent contamination.
- AI/machine learning (ML) and big data models must be devised and used to offer further intelligent decision-making and, whenever possible, should be employed directly at-the-edge for greater energy efficiency.
- Industrial IoT (IIoT) systems can provide the flexibility to tailor-make new products to help cope with ever-demanding diets.
- Remotely piloted autonomous unmanned aerial vehicles (UAVs), either flying alone or in swarms, to improve efficiency.
- Smart systems based on portable real-time pest disease diagnostics and monitoring platforms to provide rapid local and regional disease incidence alerts (georeferenced) – e.g. weather/climate information for predictive models providing risk assessments and decision support for Integrated Pest Management (IPM).
- IoT devices specialising in pests and disease measurements, such as insect traps and other systems based on image recognition or AI models.
- Large-scale and high-precision measurements of plant growth, architecture and composition.
- Winning the global platform game on various application sectors (that are currently strong) and in building effectively and, at a high level, outperforming applications and systems for industrial and business needs.
- Preparing for the 5G and beyond era in communications technology, especially its manufacturing and engineering dimension.
- Solving IoT and SoS cybersecurity and safety problems, attestation, security-by-design, as only safe, secure and trusted platforms will survive.
- Interoperability-by-design at the component, semantic and application levels.
- IoT configuration and orchestration management that allows for the (semi)autonomous deployment and operation of a large number of devices.
- Decision support for AI, modeling and analytics in the cloud and also in edge/fog settings.

In the digital society application domain, the provision of improved, robust, secure and interoperable connectivity will support the overall strategy regarding:

- Use energy and resources more efficiently within the existing installed base of industrial processes. Reduce or prevent waste.
- AI into the design, manufacturing, production and deployment processes, productivity can be improved.
- Collaborative product-service engineering, life cycle engineering: extending R&D to consider how products and systems will be integrated into the industrial service program of the company. This should possibly be enhanced by obtaining further knowledge to provide services for other similar products (competitors!) as well their own installed base.

- Remote engineering and operations, tele-presence: operating or assisting in operations of industrial systems from remote sites.
- Local and global services: organising services locally close to customers and centrally at vendors' sites.
- Edge/cloud solutions: implementing distributed service applications on effective edge-cloud systems.
- Full lifecycle tutoring: monitoring activities, level of stress and performance-oriented behavior during the product's life, from anticipating its end of life to properly handling its waste and recycling, including improved re-design for the next generation of products.

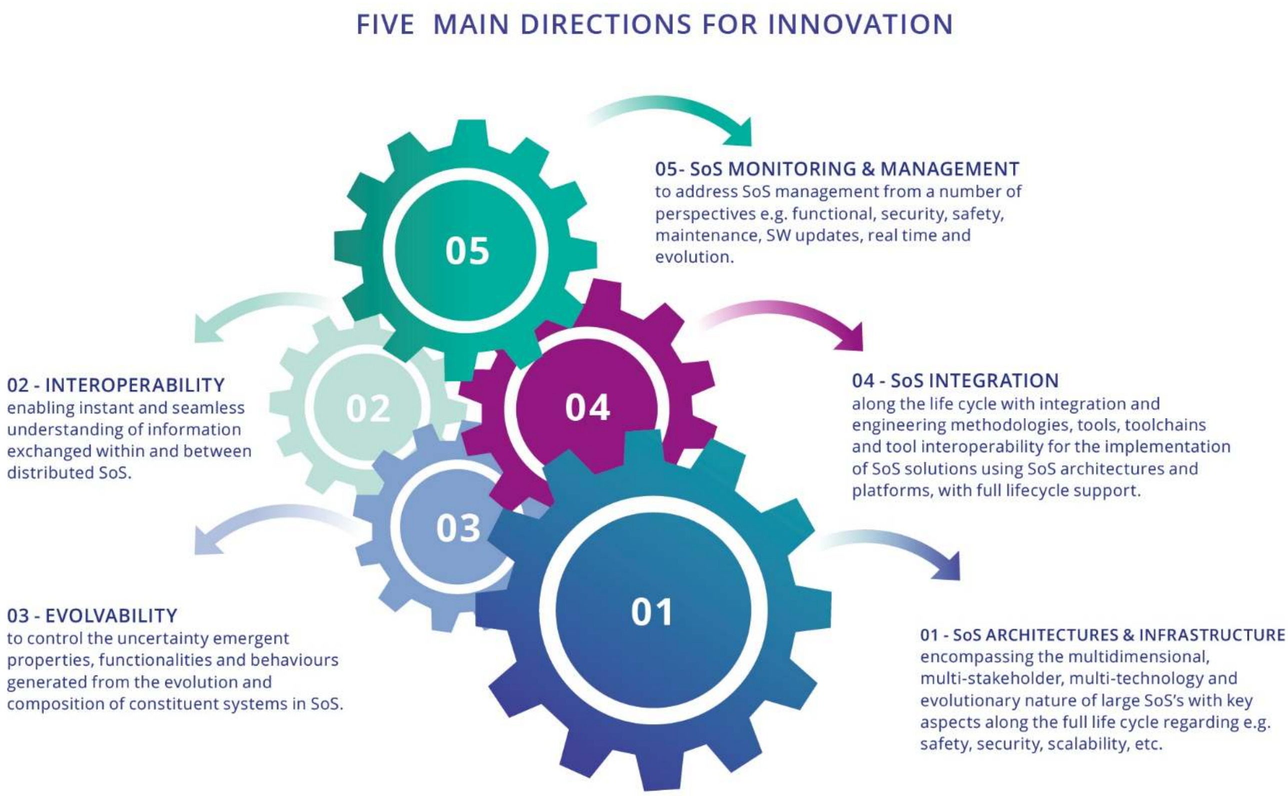


Figure 1.4.3 - Five Main Directions of Innovation (source: Eurotech).

1.4.3. MAJOR CHALLENGES

Five major challenges have been identified for the System of Systems domain:

- **Major Challenge 1**: Open SoS architecture and infrastructure.
- **Major Challenge 2**: SoS interoperability.
- **Major Challenge 3**: Evolvability of SoS composed of embedded and cyber-physical systems.
- **Major Challenge 4**: SoS integration along the life cycle.
- **Major Challenge 5**: SoS monitoring and management.

1.4.3.1. Major Challenge 1: Open SoS architecture and infrastructure

Open SoS architecture and infrastructure encompassing the multidimensional, multi-stakeholder, multi-technology and evolutionary nature of large SoS's with key aspects along the full life cycle regarding e.g. safety, security, scalability, engineering efficiency, real time performance, advanced control, QoS and distributed intelligence.

1.4.3.1.1. State of the art

SoS requires architecture and available infrastructure that encompasses the multidimensional, multi-stakeholder, multi-technology and evolutionary nature. Architecting a SoS is fundamentally different from architecting a single embedded system. The complexity of SoS architecting can be exemplified by the architecture of a complete smart city, with all its subsystem, stakeholders, technologies and evolutionary nature.

The current industrial state of the art consists in a couple of major commercial and proprietary information/communications/control/technology platforms offering industrial solutions for complex solutions from companies like e.g. Schneider Electric[3], Siemens[4], Bosch[5], Emerson[6], ABB[7], Advantech[8], AutoSAR. These proprietary digital platforms, at various levels, support design, implementation and operation of SoS architectures tailored for dedicated solutions in sectors including e.g. manufacturing, water and wastewater, minerals and mining, oil and gas, energy sectors, smart cities and automotive. It is also clear that the fundamental computer science basis for these products is quite old.

The current industrial state-of-the-art SoS's are based on extensions to existing major enterprise resource planning (ERP), manufacturing execution system (MES), supervisory control and data acquisition (SCADA), distributed control systems (DCS), robot controllers (RC), computer numerical controllers (CNC), and programmable logic controllers (PLC) products. Such extensions are mostly based on a central service bus concept. Such service buses are responsible for integrating legacy ERP, MES, SCADA, DCS, RC, CNC and PLC technologies from multiple vendors, at best. For emerging SoS application areas like autonomous driving, smart energy grid, smart agriculture and smart cities, the SoS technology is still in an emerging phase. Still Europe is the leading player for industrial automation and digitalisation, with a very strong position in the upcoming areas of autonomous driving, smart energy, smart agriculture and smart cities.

To take the next step, Europe and other regions have invested in a number of open SoS integration frameworks and platforms. A summary of these is shown in Figure [9].

---

[3] https://ecostruxure.schneider-electric.com/
[4] https://www.plm.automation.siemens.com/ global/en/webinar/iiot-the-next-big-digital- disruption/31921
[5] https://blog.bosch-si.com/bosch-iot-suite/
[6] https://www.emerson.com/de-de/ automation/operations-business- management/plantweb-digital-ecosystem
[7] https://ability.abb.com/
[8] https://www.advantech.com/resources/ news/advantech-launches-30-iiot-solutions- through-the-co-creation-model-and-the- wise-paas-platform-and-announces-a- large-scale-showcase-in-november
[9] Industrial Frameworks for Internet of Things: A Survey, IEEE System journal 2020

Most platform initiatives are based on Service Oriented Architectures (SoA) and microservices, which points towards a primary technology for such platforms. Although none of these open SoS platforms are currently in wide commercial usage, early examples can be found in small IoT solutions in various application areas. Major industrial usage remains rare, but MES-level adoption can be found in automotive production, for example.

Open architectures and reference implementations such as e.g. the IMC-AESOP approach[10], Eclipse Arrowhead[11], Eclipse Basyx[12], FiWare[13], PERFoRM30[14] are providing a link to standardisation activities in national and international innovation platforms. In the automotive domain, AutoSAR is developing in the microservice direction. Such standardisation activities are e.g. DIN Specification 91345[15] "Reference Architecture Model for Industry 4.0" (RAMI 4.0), the "Industrial Internet Architecture" (IIA), the "High Level Architecture of the Alliance for Internet of Things Innovation", the "NIST Big data Reference Architecture", to name just a few.

Europe has strongly invested in large projects that have delivered open platforms for the implementation of solutions-based on SoS platforms[16]. Considering the platforms referred to in 1.4.4, Eclipse Arrowhead, AUTOSAR, FiWare and BaSyx have all been developed with substantial European leadership and partnership.

---

[10] https://link.springer.com/ book/10.1007/978-3-319-05624-1

[11] https://www.taylorfrancis.com/ books/e/9781315367897

[12] https://www.eclipse.org/basyx/

[13] https://www.fiware.org/

[14] https://www.taylorfrancis.com/ books/e/9780429263316

[15] https://www.en-standard.eu/din-spec-91345-reference-architecture-model-industrie-4-0-rami4-0/

[16] From Internet of Things to System of Systems – Market analysis, achievements, positioning and future vision of the ECS community on IoT and SoS, P Azzoni, Artemis 2020.

FEATURES	ARROWHEAD	AUTOSAR	BASYX
Key principles	SOA, local automation clouds	Run-time, electronic control unit (ECU)	Variability of production processes
Realtime	Yes	Yes	No
Run-time	Dynamic orchestration and authorisation, monitoring, and dynamic automation	Run-time environment (RTE) layer	Run-time environment
Distribution	Distributed	Centralise	Centralise
Open source	Yes	No	Yes
Resource accessibility	High	Low	Very low
Supporters	Arrowhead	AUTOSAR	Basys 4.0
Message patterns	Req/Repl, Pub/sub	Req/Repl, Pub/sub	Req/Repl,
Transport protocols	TCP, UDP, DTLS/TLS	TCP, UDP, TLS	TCP
Communication protocols	HTTP, CoAP, MQTT, OPC-UA	HTTP	HTTP, OPC-UA
Third-party and legacy systems adaptability	Yes	Yes	Yes
Security manager	Authentication, authorisation and accounting Core system	Crypto service manager, secure onboard communication	--
Standardisation	Use of existing standards	AUTOSAR standards	Use of existing standards

FIWARE	IOTIVITY	LWM2M	OCFW
Context awareness	Device-to-device communication	M2M, constrained networks	Resource-oriented REST, Certification
No	Yes (IoTivity constrained)	No	No
Monitoring, dynamic service selection and verification	No	No	No
Centralise	Centralise	Centralise	Centralise
Yes	Yes	Yes	No
High	Medium	Medium	Low
FIWARE Foundation	Open Connectivity Foundation	OMA SpecWorks	Open Connectivity Foundation
Req/Repl, Pub/sub	Req/Repl, Pub/sub	Req/Repl	Req/Repl
TCP, UDP, DTLS/TLS	TCP, UDP, DTLS/TLS	TCP, UDP, DTLS/TLS, SMS	TCP, UDP, DTLS/TLS, BLE
HTTP, RTPS	HTTP, CoAP	CoAP	HTTP, CoAP
Yes	No	No	No
Identity manager enabler	Secure resource manager	OSCORE	Secure resource manager
FIWARE NGSI	OCF standards	Use of existing standards	OCF standards

Figure 1.4.4 - Open SoS integration frameworks and platforms[17]

For the cross-domain requirements on e.g. security, safety, evolution application and business critical details need to be considered. As an example thereof security takes on new dimensions in the case of SoS. In this Chapter, security is taken to be the ability to prevent leaking information and to prevent the taking over of control of the SoS by agents not being part of the SoS, but also the guarantee that no hostile party can prevent the sharing of essential information between the systems comprising the SoS. Several security aspects require attention. First, the level of security of each individual system requires attention: the lower bound to security of an SoS is determined by the system with the lowest security level, and by the link with the lowest security level between systems ("weakest link in the chain"). Thus, requirements like Quality, Reliability, Safety and Cybersecurity at the system and SoS levels are covered in Chapter 2.4 of this ECS-SRIA.

[17] Industrial Frameworks for Internet of Things: A Survey, C. Paniagua and J. Delsing, in IEEE Systems Journal, vol. 15, no. 1, pp. 1149-1159, March 2021, doi: 10.1109/JSYST.2020.2993323.,

However, combining a very large number of systems in an SoS can result in a lower overall security level than the lowest security level of any system in the SoS: an attacker can now combine and relate information from two or more systems, which in combination can reveal new information. Segmentation of an SoS is thus of large importance both for architecture but also for actual implementation and maintenance and updates of the life cycle.

Systems must not only defend against and monitor possible attacks, but also measures must be taken avoiding infection by intrusions from one system to the other systems in the SoS. Only this way resilience and cybersecurity can be attained.

The spectrum of systems making up an SoS includes both systems in the cloud, where security can be closely monitored as in e.g. data warehouses, and systems at the edge. Edge systems pose a higher level of cyber insecurity because of the limited resources often available at the edge (e.g. power, communication bandwidth).

Another aspect is SoS safety. Here architectures and platforms need to address safety from various application domains and their respective standards and regulations. More details related to the ECS application domain requirements on Quality, Reliability, Safety and Cybersecurity at the system and SoS level are covered in Chapter 2.4 of this ECS-SRIA.

1.4.3.1.2. Vision and expected outcome

This Major Challenge is expected to lead to a set of EU-strategic open SoS architectures and infrastructures. From such infrastructures, vendor and large company platforms can be devised, capable of supporting a wide range of solutions in diverse fields of applications covering the ECS supply chain and supporting efficient life cycle management.
This requires new and improved infrastructure technologies comprising:

- Robust design- and run-time infrastructure enabling integration and orchestration of functionalities from edge to cloud.
- Infrastructure support for multi-level security, security management, safety, safety management, scalability, engineering efficiency, real-time performance, closed loop and digital control, QoS, distributed intelligence and other key application area requirements.
- Interoperability to legacy SoS technology (“to-the-past”).
- Interoperability to existing and emerging IoT and SoS technologies and infrastructures (“to-the-future”).
- Support for autonomous operation, resilience, fail-over and mitigation management.
- Enabling SoS flexibility.
- Engineering support through model based engineering and associated domain specific languages (c.f. Chapter 2.3 Architecture and Design: Method and Tools).

The expected outcome is a set of EU-strategic open source platforms. These infrastructure platforms should have long-term governance with industry-friendly licensing schemes such as e.g. Eclipse ECL2. Such platforms should also have strong EU-based value chain support.

To cope with increasing complexity, the SoS engineering community is constantly researching improvements to its engineering processes. To ensure the complexity remains manageable, modeling approaches are used. The challenge in these approaches is to find the right level of abstraction that also allows for reasoning about the system while still containing sufficient information to connect to lower levels of abstraction, often by generating code for some underlying implementation platform.

It is not only that the complexity of the SoS is growing, but there are also extra-functional requirements that are often interlinked playing an increasingly important role. For example, with the demand for greater speed and the concomitant energy consumption, systems are often required to process information quickly but within a tight energy budget. These two requirements are clearly conflicting and choosing the right trade-off can be a balancing task. With the realisation that the planet's resources are limited, as exemplified in the European Green Deal, also comes the demand for resource conservation, resulting in more and intertwined requirements, putting greater demand on the dynamic and evolution capabilities of both the SoS architectures and the architecture tools that support the complexity of SoS.

Some important but necessary aspects of SoS architecture are:
- Security and trust,
- Safety,
- Robustness,
- Composability,
- Evolution,
- Interoperability (data exchange and data models),
- Engineering tools and procedures,
- Energy consumption,
- Unified environmental data model,
- Environmental footprint optimisation,
- Resilience.

1.4.3.1.3. Key focus areas

The key focus is how SoS architectures and their open infrastructure can enable and leverage important and necessary aspects while also enabling efficient adaptation to specific application solutions.
To support EU strategic autonomy, a small number of SoS architectures and integration platforms should be driven by EU-based ecosystems. Important features that such platforms should provide include:
- Robust SoS infrastructure capable of supporting a wide range of solutions in diverse fields of applications,
- SoS infrastructure and associated engineering tools and toolchains that support the complete engineering processes in both design- and run-time, including SoS critical aspects such as e.g. security, safety and risk mitigation,
- Suitable and adaptable engineering processes, with associated training material for solution engineering.

- Methods for the handling of (often wide-spread) legacy elements, e.g. as black box models.

2.4.3.1. Major Challenge 2: SoS interoperability

SoS interoperability enables instant and seamless understanding of information exchanged within and between networked and distributed systems.

2.4.3.1.1. State of the art

Interoperability in the SoS domain is a rising problem for cost-effective engineering and operation of systems of embedded and cyber-physical systems (see Figure 1.4.).

There is currently no industrial solution to this problem. Academia and industry are experimenting with approaches based on, for example, ontologies[18], machine learning[19], model-based engineering and open semantic frameworks[20]. Even if no clear winning approach can be identified based on current research results, growing interest can be noted for e.g. ontology, data and model driven approaches. Automating considerable parts of interoperability engineering (design-time and run-time) will improve SoS operational quality and will be very cost efficient.

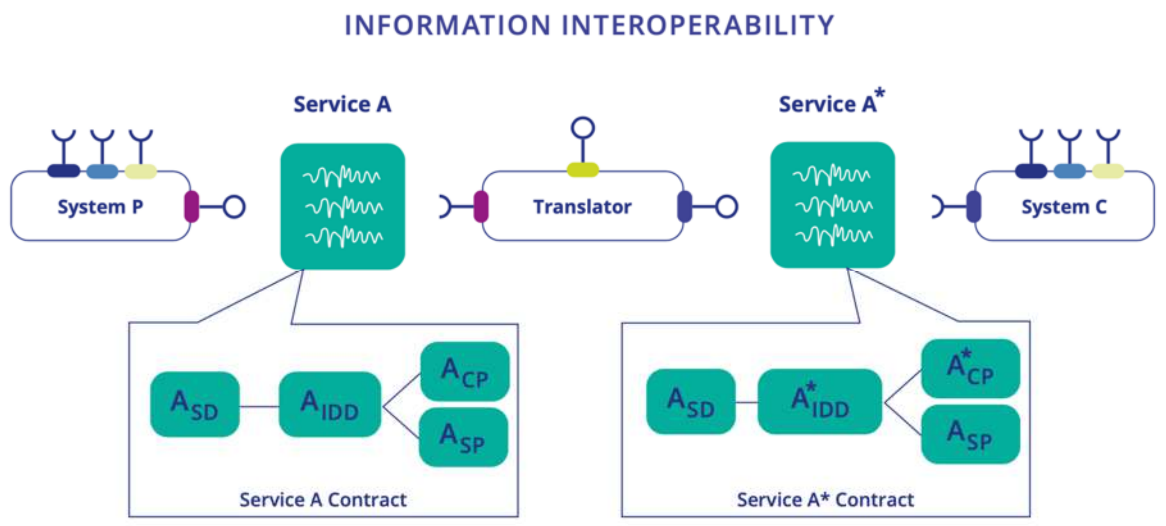


Figure 1.4.5 - Information interoperability between two service providers can be addressed by means of translators. The design of such translators for the payload information is currently necessary to provide for every situation where interoperability is requested.

[18] Extended semantic annotations for generating translators in the arrowhead framework, F Moutinho, L Paiva, J Köpke, P Maló - IEEE Transactions on Industrial Informatics, 2017
[19] Interoperability and machine-to-machine translation model with mappings to machine learning tasks, Jacob Nilsson, Fredrik Sandin and Jerker Delsing, IEEE INDIN 2019
[20] An open semantic framework for the industrial Internet of Things, S Mayer, J Hodges, D Yu, M Kritzler, F Michahelles - IEEE Intelligent Systems, 2017

2.4.3.1.2. Vision and expected outcome

To enhance EU leadership and sovereignty in the field of SoS based on embedded and cyber-physical systems, autonomous information translation for understanding is a necessity. Some integration platforms already focus on protocol and information interoperability (Derhamy, 2018[21]). To enable the cost- and time-efficient engineering of solution integration and extension, their updates and upgrades over the lifecycle is crucial. Therefore, SoS integration platforms have to provide mechanisms for dynamic and instant information translation across the ontologies and semantics used the individual constituent systems of the SoS.

2.4.3.1.3. Key focus areas

To facilitate substantial cost reductions for SoS solutions, autonomous and dynamic mechanisms for information translation are required. Such mechanisms should cover:

- Translation between standardised data models (e.g. ISO 10303[22], ISO 15926[23], BIM[24]).
- Translation between different implementations of standardised data models.
- Automated data model translation.
- Autonomous data model translation.
- Efficient and flexible engineering procedures.
- Engineering tools that support the complete engineering process in both design- and run-time.
- Support for key automation requirements.
- Automated translation engineering e.g. AI-driven, model based code generation.

3.4.3.1. Major Challenge 3: Evolvability of SoS composed of embedded and cyber-physical systems

SoS intrinsic nature is dynamic and SoS evolve with components, functions and purposes added, removed, and modified along their continuously evolving lifecycle (a life cycle that potentially never finishes). An SoS has properties, behaviours and functionalities that mainly do not reside in any constituent system but in the SoS as a whole and allow the SoS to achieve its own goals. These properties, functionalities and behaviours at the SoS level emerge in a direct relationship to the SoS evolution and, being potentially unknown, must be monitored and managed, i.e., detected, identified, understood and controlled. Because the results of the composition/evolution could be uncertain, SoS architectures and platforms, open and proprietary in conjunction with the proper engineering support

[21] H. Derhamy, J. Eliasson and J. Delsing, "IoT Interoperability—On-Demand and Low Latency Transparent Multiprotocol Translator," in *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1754-1763, Oct. 2017, doi: 10.1109/JIOT.2017.2697718.

[22] https://www.iso.org/standard/66654.html

[23] https://15926.org/home/

[24] https://en.wikipedia.org/wiki/Building_information_modeling

(methods and tools), should provide solutions to manage the evolution and resulting uncertainty emergent properties, functionalities and behaviours.

3.4.3.1.1. State of the art

Evolvability and composability are multi-dimensional aspects of SoS evolution, that affect SoS architectures, properties, functionalities and behaviours from different perspectives (evolvability, trust, interoperability, scalability, availability, resilience to failures, etc.). Primarily, composability must ensure the persistence of the five major attributes that characterise an SoS (see Maier, 1998[25]). Vertical (hierarchical) composability provides the most common way to build an SoS that is typically structured in a hierarchical stack composed of adjacent layers. Vertical composability has to deal with the different abstraction levels of the stack layers, adopting aggregation and de-aggregation solutions as references to compose the constituent systems of the SoS. Architectural composability, on the other hand, is fundamental for SoS design, specifically when critical requirements such as trust or safety must be satisfied (see Neumann 2004[26], for an extensive report on trustworthy composable architectures).

In the hierarchical structure of an SoS, the constituent systems that are at the same level typically compose horizontally (in parallel or serially), potentially generating competing chains of constituent systems. Serial composability represents a critical issue for all properties that are not automatically transitive, such as trust. Indeed, the inclusion of AI in embedded and cyber-physical systems increases the required level of trust, as well as the uncertainty of the results of the composition process (see, for example, Wagner, 2015[27]).

When the constituent systems expose high-level services, service composability allows for the creation and provision of new added-value services at the SoS level, combining the resources, functionalities, information, etc., of the constituent systems. Eventually, the engineering process deals with composability, enabling it by design (already present from the constituent systems level) and/or managing it during the operations of the SoS, to address the dynamic nature of SoS in time (run-time composability associated with evolutionary development and potential emergent properties, behaviours, and functionalities).

3.4.3.1.2. Vision and expected outcome

The dynamic nature of SoS is based on the composition and integration of embedded and cyber-physical systems. The role of composability is to ensure that functional and extra-functional properties (scalability, quality of service (QoS), performance, reliability, flexibility, etc.), and the functionalities and behaviours of the constituent systems are preserved in the

---

[25] Architecting Principles for Systems-of-Systems, Mark W. Maier, Systems Engineering journal, John Wiley & Sons 1998

[26] Peter G. Neumann, "Principled Assuredly Trustworthy Composable Architectures", DARPA, Computer Science Laboratory SRI International EL-243, 333 Ravenswood Ave, Menlo Park, California 94025-3493, USA.

[27] Wagner, M.; Koopman, P. A Philosophy for Developing Trust in Self-driving cars. In Road Vehicle Automation 2; Meyer, G., Beiker, S., Eds.; Lecture Notes in Mobility; Springer: Cham, Switzerland, 2015; pp. 163–171, doi:10.1007/978-3-319-19078-5_14.

SoS or combined in a predictable and controlled way, even when the constituent systems recombine dynamically at run time. The lack of solutions to dynamically manage composability represents one of the limitations hindering the market uptake and diffusion of SoS.

Composability should be conceived as a quality of SoS that makes them future proof: (i) the relationships between components that allow them to recombine and assemble in different and potentially unlimited architectural combinations, and ensure and exploit the re-use of components; (ii) the extension of components lifetime within the evolution of the SoS during its lifecycle; (iii) the possibility that SoS will easily evolve, adapting to new contexts, new requirements and new objectives; and (iv) the simple substitution of faulty, inadequate and/or new components with a minimal impact for the SoS, guaranteeing the survival and sustainable evolution of the SoS. Composability also has to consider cross-sectorial requirements like e.g. security, safety, trust, evolution.

Ensuring composability at the SoS level represents a very challenging goal, potentially generating serious and critical consequences, and even preventing the integration of the SoS. Indeed, considering a property that characterises a constituent system with a certain attribute, it is not guaranteed that the same property will characterise it when the constituent system becomes part of an SoS. In addition, if the property is still present, it is not guaranteed that it will have the same attribute. The same applies to the constituent system's functionalities, behaviours, etc.

As a consequence, one major effect of the composition, integration, evolution of the constituent systems is the evolution of the SoS, with emergent properties, functionalities and behaviours which generate uncertainty. For instance, when SoS evolution affects security, safety, trust, interoperability, scalability, availability, resilience to failures, etc., the impact of the uncertainty could potentially be extremely serious.

The inclusion of AI in SoS increases the importance of composability, because it may significantly increase the complexity, variability and fuzziness of composability results. AI enables a completely new category of applications for SoS. Therefore, the availability of specific solutions for the validation, verification and certification of SoS composed of AI-based systems is a critical requirement.

Predicting and controlling the effects of composability is also fundamental for the interaction of humans along the SoS lifecycle and the protection of human life should be ensured in SoS evolution. Uncontrolled and unmonitored composition could lead to deviations from expected behaviours or generate unknown emergent behaviours potentially dangerous for humans. The increasing level of automation introduced by SoS accentuates this criticality, and will require that humans still intervene in cases of emergency (for example, in automated driving).

The solutions proposed to manage composability will also have to support the multi-domain nature of SoS, the presence of different stakeholders in its lifecycle, and the different regulations and standards that apply to these domains. From an engineering perspective, emergent behaviours require that the development of SoS, applying composability, is evolutionary and adaptive over the SoS continuously evolving lifecycle, which potentially may never finish. In fact, SoS architectures and platforms, jointly with the proper engineering support, will have to provide solutions to control the uncertainty of evolvability and ensure adequate countermeasures.

3.4.3.1.3. Key focus areas

Since the technology base, and the organisational and human needs are changing along the SoS lifecycle, SoS architecting will become an evolutionary process based on composability. This means: (i) components, structures, functions and purposes can be added; (ii) components, structures, functions and purposes can be removed; or (iii) components, structures, functions and purposes can be modified as owners of the SoS experience and use the system. In this sense, the dynamically changing environmental and operational conditions of SoS require new architectures that address the SoS goal(s), but thanks to composability will also evolve to new system architectures as the goal(s) change.

Evolution in SoS is still an open research topic requiring significant effort and the key areas of research and innovation include:

- Methods and tools for engineering evolvability of systems of embedded and cyber-physical systems, e.g. AI driven, model based (c.f. Chapter 2.3. Architecture and Design: Methods and Tools).
- Evolutionary architectures in systems of embedded and cyber-physical systems.
- Evolvable solutions for trust, availability, scalability, and interoperability.
- Evolvable solutions capable for managing resulting uncertainty emergent properties, functionalities and behaviours, including resilience to failures.
- Evolvability in systems of cyber-physical systems through virtualisation, e.g. digital twins.
- Methods and tools to manage emergencies in embedded and composable systems of cyber-physical systems.
- Service-based vertical and horizontal evolvability to enable high-level, and potentially cross-domain, interoperability of embedded and cyber-physical systems.

4.4.3.1. Major Challenge 4: SoS integration engineering along the life cycle

Integration and engineering methodologies, tools, tool chains and tool interoperability are fundamental to enable the implementation of SoS solutions using SoS architectures and platform technologies, supporting the whole lifecycle.

4.4.3.1.1. State of the art

Europe is a world leader in the engineering of systems of systems. Major European companies such as Siemens, ABB, Schneider, Valmet, Bosch and Endress+Hauser, together with a number of large system integration companies (e.g. Afry, VPS and Midroc), offer complete engineered solutions, making Europe the leading global automation SoS provider.

Most solutions for embedded and cyber-physical systems engineering are based on highly experienced teams of engineers supported by a heterogeneous set of SoS engineering tools. For example, engineering practice and associated standards provide design-time solutions based on, for example, IEC 61512 (ISA 88)[28], IEC 62264 (ISA95)[29], IEC81346[30], ISO 10303, ISO 15924, IEC 62890[31]. The proposed Industry 4.0 architectures, formally provided by the DIN specification 91345 RAMI 4.0, have not yet made it into industrialised engineering procedures, or associated tools and toolchains. Many of these standards investigate updates of their data models to be based on e.g. ontologies and semantic web.

The current state of the art engineering of SoS remains more an art than a well-structured integration and engineering process. For example, the analysis of emergent behaviour of very large SoS is still at a foundational research level in academia.

4.4.3.1.2. Vision and expected outcome

The European leadership in application fields such as distributed automotive and industrial automation and digitalisation indicates some excellent skill sets in the art of SoS engineering. In the short to medium term, Europe has to transfer these skills into systematic and robust engineering procedures supported by integrated and efficient tools and tool chains. Please also refer to Chapter 2.3 and Chapter 1.3

This is expected to lead to engineering processes, tools and tool chains covering the whole life cycle that to significant extent can be automated while supporting integration between multiple stakeholders, multiple brand and multiple technologies. To support such integration and engineering efficiency, solution quality and sustainability concrete advancements like in Figure 1.4.6[32] will become necessary. The advancement may include integration and engineering process capabilities like:

- Flexible integration and engineering procedures.
- Model-based engineering procedures and tool,
- Supported by interoperable and flexible toolchains.
- Integration of multi-stakeholder engineering processes.
- Automation of substantial parts of the integration and engineering process.

---

[28] https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa88
[29] https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa95
[30] https://en.wikipedia.org/wiki/IEC_81346
[31] https://webstore.iec.ch/preview/info_iec62890%7Bed1.0%7Den.pdf
[32] Urgese, G.; Azzoni, P.; van Deventer, J.; Delsing, J.; Macii, A.; Macii, E. A SOA-Based Engineering Process Model for the Life Cycle Management of System-of-Systems in Industry 4.0. *Appl. Sci.* 2022, *12*, 7730. https://doi.org/10.3390/app12157730

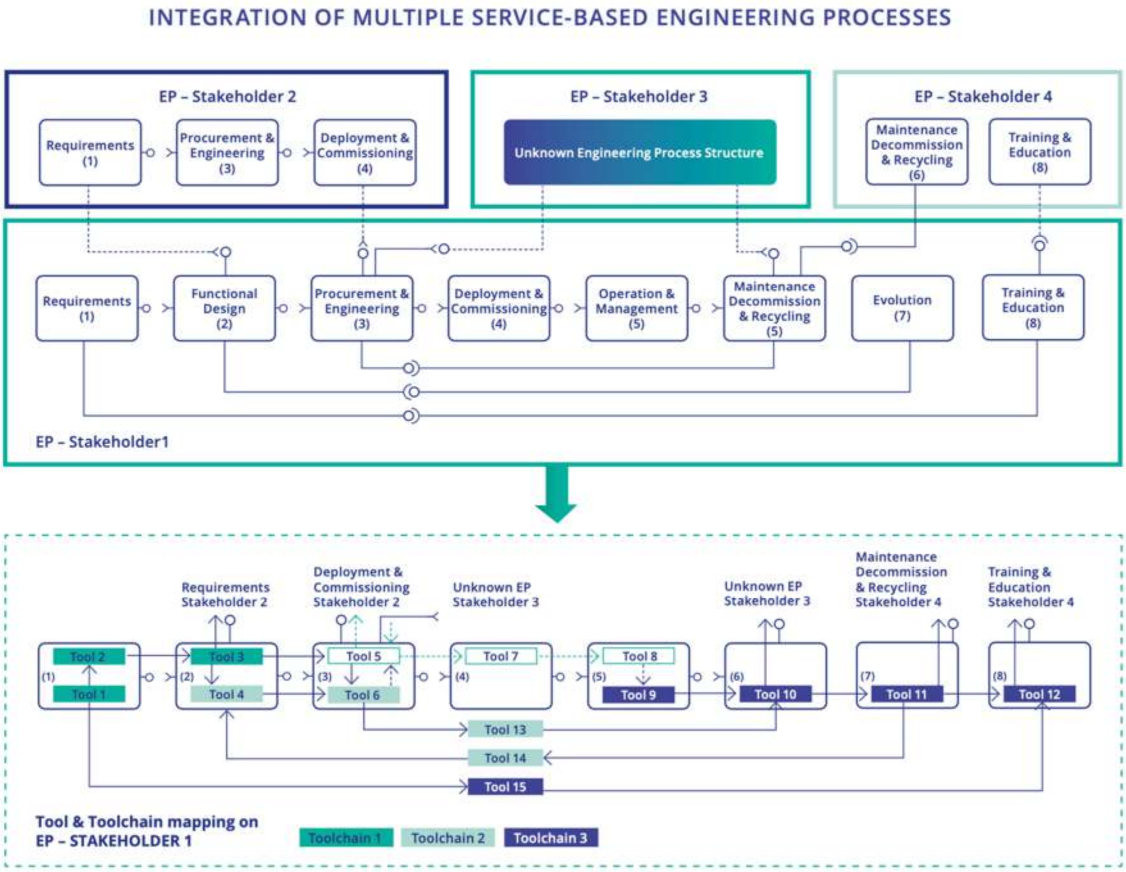


Figure 1.4.6 - Example of conceptual service-oriented view on the integration of multiple service-based engineering processes (EP) from different stakeholders, including the engineering process mapping with integrated toolchains and tools.

4.4.3.1.3. Key focus areas

In support of EU leadership and sovereignty in the field of SoS engineering the ambition is to invest in a small number of architecture infrastructures and their associated tools, tool chains and engineering processes. Strong European-based ecosystems should be created and provided with long-term governance also connected to open source. These engineering processes, methodologies, tools and toolchains shall provide, for example:

- Efficient, flexible and automated engineering processes.
- Model-based engineering.
- AI-supported engineering tools and processes
- Engineering tools supporting the complete engineering process along the system's lifecycle.
- Support for key automation requirements.
- Automated engineering, low-code engineering.
- SoS traceability and analytics interoperable with engineering tools and tool chains.
- SoS evolution impact analysis.
- Automated testing validation and verification (TV&V) along the life cycle.

In particular, SoS TV&V introduces a significant challenge, mainly due to complexity, to the effects of composition (not always known in advance) and to SoS dynamic evolution over time. For SoS, a full TV&V procedure prior to deployment is practically unrealistic. Typically, the TV&V of each constituent system is asynchronous and independent of SoS, challenging the SoS TV&V with feature and capability evolution. For this motivation, a structured framework methodology and tools is necessary to demonstrate an appropriate level of confidence that the feature under test is present in the SoS, and that no undesirable behaviours are also present. This implies a need for end-to-end system capabilities metrics and, according to the flow of data, control and functionalities across the SoS, additional test points, recurring tests and AI-empowered data collection. This analysis should be considered to address changes in the constituent systems and to receive feedback on anomaly behaviours.

5.4.3.1. Major Challenge 5: SoS monitoring and management

Management of SoA-based SoS will require structured and scalable approaches to status monitoring and strategies and methodologies to address SoS management from a number of perspectives e.g. functional, security, safety, maintenance, SW updates, real time and evolution. It is clear that a high degree of automation and autonomy need to be introduced to keep quality up and cost down. Management and monitoring of SoS play a particularly crucial role when application faults result in personal injuries or property and environmental damages: e. g. critical infrastructures (electric grid, rail network) connected (semi)-autonomous automotive systems, medical monitoring, industrial plant control systems, robotics and automatic pilot avionics.

5.4.3.1.1. State of the art

Current industrial state of the art for monitoring and management of SoS reflects back to monitoring and management of production automation, energy grid automation and similar. Looking closer we find a plethora of commercial application solutions tailored to specific applications. Many of these are very application and site specific and “home brewed”.

There is a wide set of different realms to be monitored and managed, ranging from modern production processes, smart grids, smart cities, automotive traffic networks, only to name some of them. Furthermore, for each of these realms their operation requires different competences and groups within an organization, and it follows different guidelines. Some examples are:

- Status of operation
- Safety
  Real time performance
- Real time monitoring of sensors and actuators, incl. fault detection and isolation
  Validation of signals (using redundancies created by the data network of the SoS)
- Control
- Maintenance
- Assets
- Security

These aspects do have more or less known and understood relationships/dependencies which also will change in run-time. This provides a monitoring and management landscape which is very heterogeneous and dynamic. As a result, management methodologies need to be supported by automated and autonomous control technology. In this context and in view of limiting data traffic in an SoS, synchronisation of systems becomes a major goal as it is directly linked to the stability of the management system.

In addition, management of complex cyber-physical SoSs must address scalability (i.e. to deal with a variable number and interconnection of systems and automated control loops) and network phenomena (such as computation/communication latency, data loss). Looking at the aspect of data management, open SoS control platforms should ensure information security management, SoS scalability, SoS engineering efficiency and also SoS real-time performance.

A wide set of tools is available, each supporting one or a few of these dimensions. In most cases these tools mandate underlying information sources and data models, which sometimes correlates with current major industrial standards like ISA95, BIM, ISO 15926 and ISO 10303.

In summary a very complex and heterogeneous landscape of, to a large extent non-interoperable, tools and methodologies with no or little capacity to be integrated across SoS dimensions.

5.4.3.1.2. Vision and expected outcome

The emerging closer digital integration of industrial and societal functionalities and domains requires SoS integration and associated monitoring and management in very complex and heterogeneous environments. The current state of the art is far from efficiently enabling this. Such enabling will require closer cooperation and integration between several levels of the ECS domain stack and society policies and governance. An example thereof is the integration and functional interoperability between open and proprietary SoS architecture and implementation platforms which reside under different jurisdictions. Here solution requirements on lifecycle and evolution as well need to be considered.

5.4.3.1.3. Key focus areas

To advance towards the vision technology and knowledge steps are required regarding:

- Monitoring and management strategies and architectural concepts in OT-IT environments.
- Methodologies and technologies for monitoring and management of multiple and interrelated SoS dimensions.
- Processes and technology for life cycle monitoring and management over SoS dimensions and society borders.
- Engineering support, tools and methods, for monitoring and management strategy and policy implementation
- Tools for control system analysis of SoS.

- Considering humans, environment and the economy in the loop.
- Engineering tools and methods for SoS control design.
- Reduction of communication effort, variable structure, variable number of systems in control loops.
- Control system testing, validation and verification (TV&V) in design and run-time.

1.4.4. TIMELINE

The following tables illustrate the roadmaps for System of Systems.

MAJOR CHALLENGE	TOPIC	SHORT TERM (2025–2029)	MEDIUM TERM (2030–2034)	LONG TERM (2035 and beyond)
Major Challenge 1: SoS architecture and open integration platforms	**Topic 1.1**: Robust SoS integration platform capable of supporting a wide range of solutions in diverse fields of applications	Architectures and associated implementation platforms with sufficient granularity and engineering support for efficient implementation of real-world Industry 4.0 solutions	Architectures and implementation platform with support for a wide set of autonomous operation e.g. M2M business execution	Architectures with support for self-X e.g. self-healing, self-extension etc.
	Topic 1.2: integration platform and associated engineering tools and toolchains that support the complete engineering process in both design- and run-time, including SoS critical aspects such as security, safety and risk mitigation	Preliminary lifecycle support for extra-functional requirements, such as energy consumption, environmental impact that translates into maintainability, sustainability, etc.	Full lifecycle support for extra-functional requirements, such as energy consumption, environmental impact that translates into maintainability, sustainability, etc.	Autonomous management of functional and non-functional dimensions
	Topic 1.3: suitable and adaptable engineering processes with training material for solution engineering	Hardware and software tools, methodology and training material suited for training of professionals and students at university level	Model based engineering support proving partial engineering automation of solutions	Automated SW engineering for most solution engineering stages.
Major Challenge 2: SoS interoperability	**Topic 2.1**: Translation between standardised data models e.g. ISO 103030, ISO 15926, BIM, …	Translation technologies enabling translation of standardised data models and demonstrated at TRL 5-7	Fully autonomous translation	

	Topic 2.2: Translation between different implementations of standardised data models	Translation technologies enabling translation of different implementations of standardised data models and demonstrated at TRL 5-7	Full cross-domain interoperability	
	Topic 2.3: automated data model translation	Technologies and tools for automating the engineering of data model translations	Fully automated information translation	
	Topic 2.4: autonomous data model translation	Technology and tools for enabling autonomous data model translation in run-rime	Fully autonomous translation	
Major Challenge 3: Evolvability of SoS composed of embedded and cyber-physical systems	**Topic 3.1**: methods and tools for engineering evolvability of systems of embedded and cyber-physical systems	Persistence of operational independence, managerial independence, geographic distribution, emergent behavior and evolutionary development	Full predictable and controllable composition of functional and extra-functional properties	Full predictable and controllable composition of functional and extra-functional properties, also covering dynamically recombining SoS
	Topic 3.2: evolutionary architectures in systems of embedded and cyber-physical systems	Modular and evolvable architectures.	Evolvability and composability by design	Automated evolvability and composability analysis in design time and run-time
	Topic 3.3: evolvable solutions for trust, availability, scalability, and interoperability.	Modular frameworks addressing trust, availability, scalability and interoperability-	Modular frameworks and open integration platforms addressing e.g. trust, availability, scalability, interoperability	Open modular frameworks and integration platforms addressing e.g. trust, availability, scalability, interoperability, evolvability, composability
	Topic 3.4: evolvable solutions capable for managing resulting uncertainty emerging properties, functionalities and behaviours, including resilience to failures	Technology frameworks supporting self-adaptability	Failures resilience at SoS level	Automated management of uncertainty and resilience to failures.
	Topic 3.5: evolvability in SoS supported by virtual engineering (e.g. digital twins)	Virtualisation of IoT and edge services based on open SoS architectures and platforms	Automated virtualisation of IoT and edge services based on open SoS architectures and platforms	Dynamic and scalable virtualisation of IoT and edge services based for run-time optimisation on open SoS architectures and platforms
	Topic 3.6: methods and tools to manage emergencies in embedded and composable SoS.	Technology frameworks supporting emergent self-adaptability	Automated technology and tools supporting emergency self-adaptability	Autonomous technology and tools supporting emergency self-adaptability

	Topic 3.7: service-based vertical and horizontal evolvability to enable high-level, and potentially cross-domain, evolvability of SoS	Open services enabling technology and data evolvability cross-domain	Open services and integration platforms enabling technology and data evolvability cross-domain	Open services and integration platforms enabling automated technology and data evolvability cross-domain
Major Challenge 4: SoS integration along the life cycle.	**Topic 4.1**: efficient and flexible engineering processes	SoA-inspired engineering processes, toolchains and tools	Engineering support for SoS emergent behaviours	Engineering support for emergent behaviours of very large SoS
	Topic 4.2: model-based engineering	Partial automated generation of SoS software using model-based engineering and AI tools	Full automated generation of SoS software using model-based engineering	Model based engineering support providing engineering automation for very complex SoS solutions
	Topic 4.3: engineering tools supporting the complete engineering process along the system's lifecycle	Engineering tools enabling run-time engineering	Multi-stakeholders and multi-domains automated engineering process	Highly automated solution engineering in a multi-stakeholders and multi-domains SoS environment
	Topic 4.4: support for key automation requirements	SoS engineering process and tools partial support for fundamental automation requirements like e.g. real time, security, safety	SoS engineering process and tools full support for fundamental automation requirements like e.g. real time, security, safety	
	Topic 4.5: automated engineering	Automation of SoS software engineering from requirements to deployment	Technologies and tool for highly automated design time control analysis in SoS environments	Technologies and tool for autonomous run-time control analysis in SoS environments
	Topic 4.6: automated testing validation and verification (TV&V)	Automated and runtime SoS TV&V for parts of the engineering process	Automated runtime SoS TV&V for the entire engineering process	Autonomous runtime SoS TV&V
Major Challenge 5: SoS monitoring and management	**Topic 5.1**: Monitoring and management strategies and architectural concepts in OT-IT environments	Real time monitoring and management of evolving OT.IT environments	Scalable monitoring architecture applicable to large scale SoS	SoS integration platforms including scalable, and manageable monitoring capabilities
	Topic 5.2: Methodologies and technologies for monitoring and management of multiple and interrelated SoS dimensions	Functional, security and safety interrelations monitoring and management	Manageable monitoring architecture of multiple SoS dimensions	SoS management based on multi-dimensional monitoring

	Topic 5.3: Processes and technology for life cycle monitoring and management over SoS dimensions	Approaches to life cycle monitoring and management for multiple SoS dimensions. Like e.g. functionality, security and safety	SoS monitoring architecture along its life cycle	SoS integration platforms supporting SoS monitoring and management evolution along its life cycle

2

Strategic Research and Innovation Agenda 2025

CROSS-SECTIONAL TECHNOLOGIES

2.1

Cross-Sectional Technologies

EDGE COMPUTING AND EMBEDDED ARTIFICIAL INTELLIGENCE

2.1 Edge computing and embedded Artificial Intelligence

2.1.1 Summary

Edge computing and embedded AI are crucial for advancing digital technologies while addressing energy efficiency, system complexity, and sustainability. The integration of AI into edge devices offers significant benefits across various sectors, contributing to a more efficient and resilient digital infrastructure, keeping privacy by processing sensible data locally. Distributed computing forms a continuum from edge to cloud, with edge computing processing data close to its source to improve performance, reduce data transmission, latency and bandwidth, enhance safety, security and decrease global power consumption. This directly impacts the features of edge systems.

AI, especially embedded intelligence and Agentic AI, significantly influences various sectors such as productivity, environmental preservation, and transportation, enabling for example autonomous vehicles. The availability of new hardware technologies drives AI sustainability. Open-source initiatives are crucial for innovation, cost reduction, and security.

Embedded AI hardware was principally developed for perception tasks (vision, audio, signal processing) with high energy efficiency. But generative AI is also emerging at the edge; first fueled by smartphones and computers (Copilot+PC, Apple Intelligence) with the need to be able to process most of the data locally and adapting to the user's habits (fine tuning performed at the edge), and will extend into other edge applications (robotics, interfaces, high level perception of the environment). This drives new constraints not only for computing parts, but also to improve memory efficiency.

The Major Challenges are:

1. Energy Efficiency: Developing innovative hardware architectures and minimizing data movement are critical for energy-efficient computing systems. Memory is becoming an important challenge as we are moving from a computing-centric paradigm to a data-centric (driven by AI). Zero standby energy and energy proportionality to load is essential for edge devices.

2. System Complexity Management: Addressing the complexity of embedded systems through interoperability, modularity, and dynamic resource allocation in a safe and secure way. Web technologies cascade to edge (containerization, WASM, protocols, …) forming a continuum of computing resources. Using a federation of small models in a Mixture of Agents or Agentic AI instead of a very large model allows to better manage complexity and modularity while using fewer computing resources.

3. Lifespan of Devices: Enhancing hardware support for software upgradability, interoperability, and second-life applications. This will require hardware that can support future software updates, increasing memory capabilities, and communication stacks. Aggregation of various devices into a "virtual device" will allow older devices to be still useful in the pool.

4. Sustainability: Ensuring European sustainability by developing solutions aligned with ethical principles (for embedded AI) and transforming innovations into commercial successes (for example, based on open standards, such as RISC-V, and for innovative solutions such as neuromorphic computing). Europe should master all steps for new AI technologies, especially the ones based on collaboration of AI agents.

2.1.2 Scope

This chapter focuses on computing components, and more specifically on embedded architectures, edge computing devices and systems using Artificial Intelligence (AI) at the edge. These elements rely on process technology and embedded software, and have constraints on quality, reliability, safety, and security. They also rely on system composition (systems of systems) and design and tools techniques to fulfill the requirements of the various application domains.

Furthermore, this chapter focuses on the trade-off between performances and power consumption reduction, and managing complexity (including security, safety, and privacy[1]) for embedded architectures to be used in different applications areas, which will spread edge computing and AI use and their contribution to European sustainability[2].

This chapter mainly covers the elements foreseen to be used to compose AI or edge systems:

- Processors (CPU) with high energy efficiency,

---

1 Security, safety, and privacy will be covered in the chapter about "Quality, reliability, safety and security"

[2] The scope of this chapter is therefore to cover the hardware architectures and their realizations (Systems on Chip, Embedded architectures), mainly for edge and "near the user" devices such as IoT devices, cars, ICT for factories and local processing and servers (computing "on-premises"). Data centers and electronic components for data centers are not the main focus of the chapter, except when the components can be used in local processing units or local servers (local clouds, swarm, fog computing, etc.). We therefore also cover this "edge" side of the "continuum of computing" and the synergies with the cloud leading the following sections to discuss:

- innovation which is needed to transfer HPC data center capabilities to edge and/or embedded environment, and
- innovation which directly addresses the embedded market (as opposed to technologies first developed for controlled, data center like environments, then later moved to embedded domain).

The technological aspects, at system level (PCB, assembly, system architecture, etc.), and embedded and application software are not part of this chapter as they are covered in other chapters. Software is important for these programmable or configurable embedded devices but will be handled in the "embedded software" chapter. However, we will still discuss architectures that efficiently execute software ("hardware friendly" software).

In the previous editions of the SRIA, this chapter dealt on Open Source and sustainability: these are transversal topics, to be addressed in the introductory chapter of the SRIA. In particular, some of the points currently covered by this chapter 2.1 regarding sustainability could be moved to other chapters:

- How to limit CO_2 emissions incurred during ECS manufacturing is covered by chapters 1.1 and 1.2.
- Architectural approaches to limit energy consumption could be covered by chapter 2.3.

However, the RISC-V architecture and related IPs will still be covered in this chapter.

- Accelerators (for AI and for other tasks, such as security):
  - GPU (and their generic usage),
  - NPU (Neural processing unit)
  - DPU (Data processing Unit, e.g. logging and collecting information for automotive and other systems) and processing data early (decreasing the load on processors/accelerators),
  - Other accelerators xPU (FPU, IPU, TPU, XPU, …)
- Memories and associated controllers, specialized for low power and/or for processing data locally (e.g. using non-volatile memories such as PCRAM, CBRAM, MRAM for synaptic functions, and In/Near Memory Computing), etc.
- Power management.
- …[3]

In a nutshell, the main recommendation of this chapter is a paradigm shift towards distributed low power architectures/topologies, from cloud to edge (what is called the "*continuum of computing*") and using AI at the edge (but not necessarily only) leading to distributed intelligence.

2.1.3 Introduction

2.1.3.1 Positioning edge and cloud solutions

In the recent months, we saw the increasing emergence of what we can call the "continuum of computing". The "continuum of computing" is a paradigm shift that merges edge computing and cloud computing into a cohesive, synergistic system. Rather than opposing each other, these two approaches complement one another, working together to optimize computational efficiency and resource allocation. The main idea behind this continuum is to perform computation where it is most efficient. In its advanced evolution, the location of computing is not static, but dynamic allowing a smooth migration of tasks or services to where the best trade-offs, according to criteria such as latency, bandwidth, cost, processing power and energy, are available. It also permits to create "virtual meta devices" by interconnecting

---

[3] Of course, all the elements to build a SoC are also necessary, but not specifically in the scope of this chapter:

- Security infrastructure (e.g. Secure Enclave) with placeholder for customer-specific secure elements (PUF, cryptographic IPs…). Security requirements are dealt with details in the corresponding chapter. The appearance of LLMs / Generative AI calls for security measures, e.g. proof of origin/authenticity etc. will have an impact on the hardware. They should also run efficiently, in a protected environment without consuming too many resources.
- Field connectivity IPs (see connectivity chapter, but the focus here is on field connectivity) (all kinds, wired, wireless, optical), ensuring interoperability.
- Integration using chiplets and interposer interfacing units are detailed in the chapters 1.1 and 1.2.
- And all other elements such as coherent cache infrastructure for many-cores, scratchpad memories, smart DMA, NoC with on-chip interfaces at router level to connect cores (coherent), memory (cache or not) and IOs (IO coherent or not), SerDes, high speed peripherals (PCIe controllers and switches, etc.), trace and debug hardware and low/medium speed peripherals (I2C, UART, SPI etc.).

various devices with various characteristics into a more global system where each individual device (or part of a device) is executing a part of the global task. This has impact on various aspects of each device, for example increasing their lifetime by coupling them with others when they are not powerful enough to perform the requested task alone. Of course, this introduces complexity in the orchestration of the devices and the splitting of the global task into small sub-tasks, each device should be interconnected with secure protocol and authentication is key to ensure a trustable use of all devices into this federation.

An example for "continuum of computing" is Apple's approach to AI, as illustrated by the announcement of Apple Intelligence. In this system, AI tasks are performed locally on the device when the local processing resources are sufficient. This ensures rapid responses and minimizes the need for data to be transmitted over networks, which can save energy and protect user privacy. For instance, simple AI tasks like voice recognition or routine summarizing, organizing text can be handled efficiently by the device's onboard accelerators (NPU).

However, when a task exceeds the local device's processing capabilities, it is seamlessly offloaded to Apple's trusted cloud infrastructure. This hybrid approach allows for more complex computations to be performed without overwhelming the local device, ensuring a smooth user experience. Moreover, for tasks that require even more powerful AI, such as those involving large language models like Chat-GPT, the system can leverage the resources of cloud-based AI services. This layered strategy ensures that users benefit from the best possible performance and efficiency, irrespective of the complexity of the task.

The continuum of computing embodies a flexible, adaptive approach to resource utilization, ensuring that computational tasks are handled by the most appropriate platform[4]. This also has impact on the architecture of edge devices and processors, which need to be prepared to support the WEB and cloud protocols (and AI), with communication stacks, security and encryption, and containerization (executing “foreign” codes in secure sandboxes). We can observe that microcontrollers can now support the wired and wireless communication stacks (IP, Wifi, Bluetooth, Thread, …), have security IPs and encryption and can execute different OS on their different cores (e.g. a real-time OS and a Linux)[5].

[4] Some OS are designed for this kind of distribution, like HarmonyOS (https://developer.huawei.com/consumer/en/doc/harmonyos-guides-V3/harmonyos-overview-0000000000011903-V3) : « *The devices running HarmonyOS are aggregated at the system layer to form a super device, allowing flexible scaling of device hardware capabilities.*
With the support of HarmonyOS, users can integrate capabilities of their various smart devices, implementing ultra-fast connection, capability collaboration, and resource sharing among them. This way, services can be seamlessly transferred to the most suitable device, delivering smooth all-scenario experience. ». Some concepts are also developed in the open source Oniro project (https://oniroproject.org/)), also based on OpenHarmony (https://gitee.com/openharmony).

[5] The Chinese companies are pushing their low-cost MCU (often based on RISC-V architecture) with all these features, for example, the Sophgo SG2002 (https://en.sophgo.com/sophon-u/product/introduce/sg200x.html) or the Espressif ESP32-C6 (https://www.espressif.com/en/products/socs/esp32-c6).

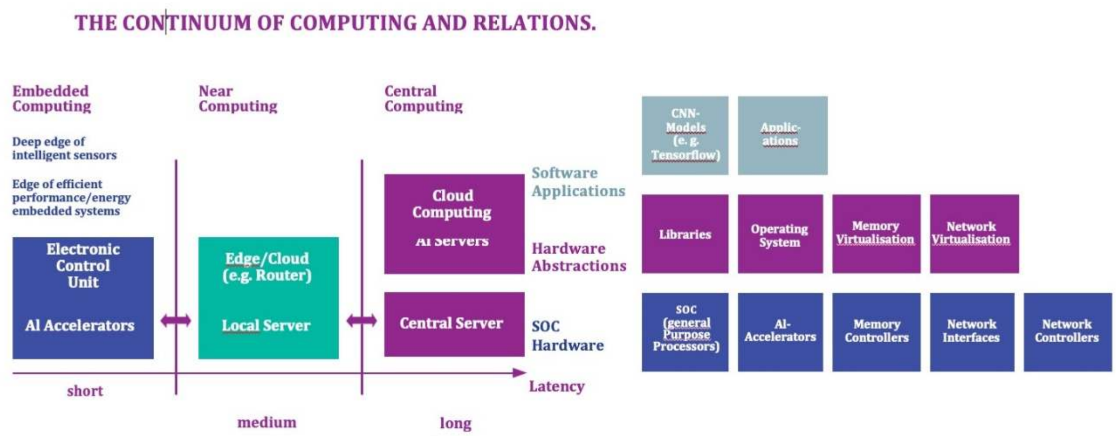


Figure 2.1.1: - The continuum of computing and relations between the elements constituting an embedded AI system (figure from Gerd Teepe)

Inside the “continuum of computing”, edge computing involves processing data locally, close to where it is generated. This approach reduces latency, enhances real-time decision-making, and can improve data privacy and security by keeping sensitive information on local devices. Cloud computing, on the other hand, offers immense computational power and storage capacity, making it ideal for tasks that require extensive resources, such as large-scale data analysis and running complex AI models.

For intelligent embedded systems, the edge computing concept is reflected in the development of edge computing levels (micro, deep, meta) that covers the computing and intelligence continuum from the sensors/actuators, processing, units, controllers, gateways, on-premises servers to the interface with multi-access, fog, and cloud computing.

A description of the micro, deep and meta edge concepts is provided in the following paragraphs (as proposed by the AIoT community).

The **micro-edge** describes intelligent sensors, machine vision, and IIoT (Industrial-IoT) devices that generate insight data and are implemented using microcontrollers built around processor architectures such as ARM Cortex M4, or recently RISC-V, which are focused on minimizing costs and power consumption. The distance from the data source measured by the sensors is minimized. The compute resources process this raw data in line and produce insight data with minimal latency. The hardware devices of the micro-edge physical sensors/actuators generate from raw data insight data and/or actuate based on physical objects by integrating AI-based elements into these devices and running AI-based techniques for inference and self-training.

Intelligent micro-edge allows IoT real-time applications to become ubiquitous and merged into the environment where various IoT devices can sense their environments and react fast and intelligently with an excellent energy-efficient gain. Integrating AI capabilities into IoT devices significantly enhances their functionality, both by introducing entirely new capabilities, and, for example, by replacing accurate algorithmic implementations of complex tasks with AI-based approximations that are better embeddable. Overall, this can improve performance, reduce latency and power consumption, and at the same time increase the devices usefulness, especially when the full power of these networked devices is harnessed – a trend called AI on edge.

The **deep-edge** comprises intelligent controllers PLCs, SCADA elements, connected machine vision embedded systems, networking equipment, gateways and computing units that aggregate data from the sensors/actuators of the IoT devices generating data. Deep edge processing resources are implemented with performant processors and microcontrollers such as Intel i-series, Atom, ARM M7+, etc., including CPUs, GPUs, TPUs, and ASICs. The system architecture, including the deep edge, depends on the envisioned functionality and deployment options considering that these devices' cores are controllers: PLCs, gateways with cognitive capabilities that can acquire, aggregate, understand, react to data, exchange, and distribute information.

The **meta-edge** integrates processing units, typically located on-premises, implemented with high-performance embedded computing units, edge machine vision systems, and edge servers (e.g. high-performance CPUs, GPUs, FPGAs, etc.) that are designed to handle compute-intensive tasks, such as processing, data analytics, AI-based functions, networking, and data storage.

This classification is closely related to the distance between the data source and the data processing, impacting overall latency. A high-level rough estimation of the communication latency and the distance from the data sources are as follows. With micro-edge the latency is below 1millisecond (ms), and the distances are from zero to max 15 meters (m). For deep-edge distances are under 1 km and latency below 2-5 ms, meta-edge shows latencies of under 10 ms and distances under 50 km, and up to 50 km (also) for fog computing. MEC concepts are combined with near-edge, with 10-20 ms latency and 100 km distance, while far-edge is 20-50ms and 200 km, and cloud and data centers are more than 50 ms and 1000 km.

	Latency	Distance
Micro-edge	Below 1ms	From 0 cm to 15 m
Deep-edge	Below 2-5 ms	Below 1km
Meta-edge	Below 10 ms	Below 50 km
Fog	10-20 ms	Up to 50 km
MEC[6] + near-edge	10-20 ms	100 km
Far-edge	20-50 ms	200 km
Cloud/data centres/HPC	More than 50 -100 ms	1000 km and beyond

[6] Multi-access edge computing (ETSI/ISG): Multi-access edge computing (MEC) brings technology resources closer to the end user. Data is processed and stored at the network's edge, not at some distant data center, significantly reducing latency.

Deployments "at the edge" can contribute, thanks to its flexibility, to be adapted to the specific needs, to provide more energy-efficient processing solutions by integrating various types of computing architectures at the edge (e.g. neuromorphic, energy-efficient microcontrollers, AI processing units), reduce data traffic, data storage and the carbon footprint. One way to reduce the energy consumption is to know which data and why it is collected, which targets are achieved and to optimize all levels of processes, both at hardware and software levels, to achieve those targets, and finally to evaluate what is consumed to process the data.

In general, the edge (in the peripheral of a global network as the Internet) includes compute, storage, and networking resources, at different levels as described above, that may be shared by several users and applications using various forms of virtualization and abstraction of the resources, including standard APIs to support interoperability.

More specifically, an edge node covers the edge computing, communication, and data analytics capabilities that make it smart/intelligent. An edge node is built around the computing units (CPUs, GPUs/FPGAs, ASICs platforms, AI accelerators/processing), communication network, storage infrastructure and the applications (workloads) that run on it.

The edge can scale to several nodes, distributed in distinct locations and the location and the identity of the access links is essential. In edge computing, all nodes can be dynamic. They are physically separated and connected to each other by using wireless/wired connections in topologies such as mesh. The edge nodes can be functioning at remote locations and operate semi-autonomously using remote management administration tools.

The edge nodes are optimized based on the energy, connectivity, size and cost, and their computing resources are constrained by these parameters. In different application cases, it is required to provide isolation of edge computing from data centers in the cloud to limit the cloud domain interference and its impact on edge services.

Finally, the edge computing concept supports a dynamic pool of distributed nodes, using communication on partially unreliable network connections while distributing the computing tasks to resource-constrained nodes across the network.

2.1.3.2 Positioning Embedded Artificial Intelligence

Even if there will be a lot of applications that will not use AI, this field is currently on top of the Gartner hype curve and more and more embedded systems will be compatible with AI requirements. We can consider the AI hardware landscape to be segmented into three categories, each defined by the processing power and application domains. These categories are Cloud AI, Embedded AI for high-performance needs, and Embedded AI for low-power applications. They also reflect the “continuum of computing” and computation should be done on all the 3 segments, where it is the most efficient according to a particular set of KPIs (such a latency, energy cost, privacy preserving, communication bandwidth, global cost, …).

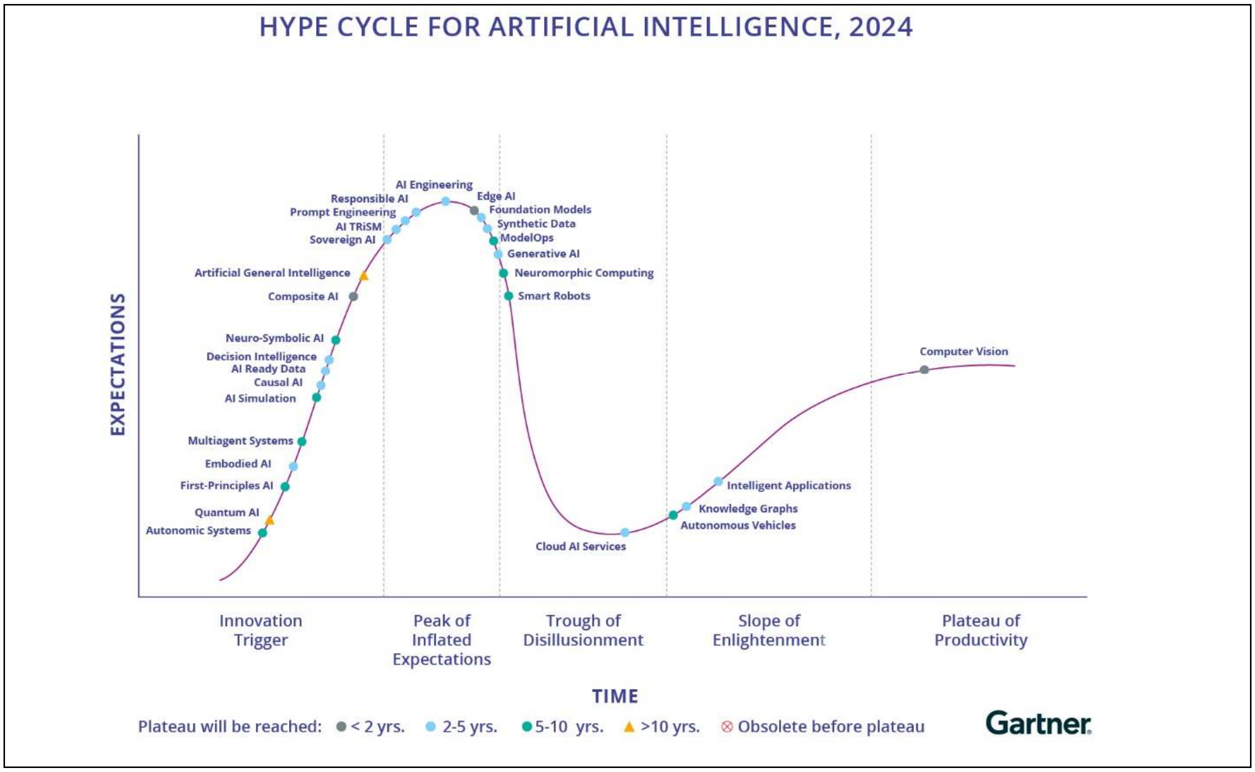


Figure 2.1.2: Hype Cycle for Emerging Technologies[7]

We can use the computing efficiency to differentiate the categories, but this is quite subjective and can change over time, but here is a classification into 3 clusters.

1. Cloud AI (~1 to 10 TOPS/W):

Cloud AI represents the highest concentration of AI processing power, leveraging GPUs capable of exceeding 2000 TOPS (Tera Operations Per Second) and large language models (LLMs) with over 4 billion parameters. This category is primarily utilized in data centers and servers, where the focus is on both training and inferencing tasks. A system (for example the Green500 system) can reach an efficiency of more than 70 GFLOPS/W in FP32 and 4 PFLOPS for the FP8 on tensor cores, or roughly 4 TFLOPS/W in FP8.

2. Embedded AI (1 to ~50 TOPS/W):

Embedded AI in the range of 1 to 100 TOPS/W focuses on bringing powerful AI capabilities closer to the source of data generation. This category employs Neural Processing Units (NPUs) with performance ranging from 40 to 2000 TOPS with multiple GB of RRAM and smaller language models (SLMs) with 1 million to 7 billion parameters. These systems are typically embedded in notebooks/PCs (for example,

---

[7] 2024 from https://www.gartner.com/doc/reprints?id=1-2HV4OPON&ct=240618&st=sb

Copilot+PC from Microsoft[8]), smartphones (for example Apple Intelligence[9]) and will be more and more used in automotive, robotics, factory automation, etc. It is also where we found the majority of "perceptive AI", used to directly analyse images, video, sound or signals, and typical AI processing methods are using Convolutional Neural Networks (CNNs) or related approaches.

For example, for automotive ADAS, Embedded AI hardware provides the necessary computational power to process sensor data in real-time, enabling advanced features like object detection, lane keeping, and adaptive cruise control. In consumer electronics such as notebooks and smartphones, Embedded AI enhances user experiences through features like voice recognition and generation, document analysis and synthesis, and intelligent photography. Additionally, in networking equipment like GPON and AP routers, Edge AI facilitates smarter data management and enhanced security measures.

3. Deep Embedded AI (>10 TOPS/W):

This category of Embedded AI is designed for applications requiring high efficiency and lower computational power. To reach high efficiency, the hardware is more specialized (ASIC), involving for example the use of Compute Near Memory (CNM) or Compute-In-Memory (CIM) architectures, and models like Convolutional Neural Networks (CNNs) or Spiking Neural Networks (SNNs) with far fewer than 4 million parameters. The processing techniques are simpler, not using Transformer-based Neural Networks, but more CNNs, Bayesian approaches etc.

The scope of this chapter is on Embedded AI and Deep Embedded AI.

2.1.4 State of the Art

The key issues to the digital world are the availability of affordable computing resources and transfer of data to the computing node with an acceptable power budget. Computing systems are morphing from classical computers with a screen and a keyboard to smart phones and to deeply embedded systems in the fabric of things. This revolution on how we now interact with machines is mainly due to the advancement in AI, more precisely of machine learning (ML), that allows machines to comprehend the world not only on the basis of various signal analysis but also on the level of cognitive sensing (vision and audio). Each computing device should be as efficient as possible and decrease the amount of energy used.

Low-power neural network accelerators will enable sensors to perform online, continuous learning and build complex information models of the world they perceive. Neuromorphic technologies such as spiking neural networks and compute-in-memory architectures are compelling choices to efficiently process and fuse streaming sensory data, especially when combined with event-based sensors. Event-based sensors, like the so-called retinomorphic cameras, are becoming extremely important, especially in the case of edge computing where energy could be a very limited resource. Major issues for edge systems, and even more for AI-embedded systems, is energy efficiency and energy management. Implementation of intelligent power/energy management policies are key for systems where AI techniques are part of

[8] https://blogs.microsoft.com/blog/2024/05/20/introducing-copilot-pcs/
[9] https://machinelearning.apple.com/research/introducing-apple-foundation-models

processing sensor data and power management policies are needed to extend the battery life of the entire system.

As extracting useful information should happen on the (extreme) edge device, personal data protection must be achieved by design, and the amount of data traffic towards the cloud and the edge-cloud can be reduced to a minimum. Such intelligent sensors not only recognize low-level features but will be able to form higher level concepts as well as require only very little (or no) training. For example, whereas digital twins currently need to be hand-crafted and built bit-for-bit, so to speak, tomorrow's smart sensor systems will build digital twins autonomously by aggregating the sensory input that flows into them.

To achieve intelligent sensors with online learning capabilities, semiconductor technologies alone will not suffice. Neuroscience and information theory will continue to discover new ways[10] of transforming sensory data into knowledge. These theoretical frameworks help model the cortical code and will play an important role towards achieving real intelligence at the extreme edge.

Compared to the previous SRIA, as explained in the previous part, we can observe two new directions:

- The recent explosion of AI, and more precisely of Generative AI drove development of new solutions for embedded systems (for now mainly PCs and smartphones), which a potentially very large market growth.
- RISC-V is coming to various ranges of products, and especially from ICs from China that support many features for connectivity (security), IOs and accelerators. Most advanced MCUs also have a small core, mostly always-on, in charge to wake-up the rest of the system in case of an event.

AI accelerator chips for embedded market have traditionally been designed to support convolutional neural networks (CNNs), which are particularly effective for image, audio, and signal analyses. CNNs excel at recognizing patterns and features within data, making them the backbone of applications such as facial recognition, speech recognition, and various forms of real-time data and signal processing[11]. These tasks require significant computational power and efficiency, which dedicated AI chips have been able to provide with new and specialized architectures (embedded NPU). **Most accelerators available today in edge computing systems are used and designed for perception applications and further developments are required to obtain further gains in efficiency.**

Generative AI at the edge (mainly for Embedded AI): the rise of (federation of) smaller models

New deep learning models are introduced at an increasing rate and one of the recent ones, with large applications potential, are transformers, which are the basis of LLMs. Based on the attention model[12], it is a "sequence-to-sequence architecture" that transforms a given sequence of elements into another sequence. Initially used for NLP (Natural Language Processing), where it can translate one sequence in a first language into another one, or complement the beginning of a text with potential follow-up, it is now extended to other domains such as video processing or elaborating a sequence of logical steps for robots. It is also a self-supervised approach: for learning it does not need labelled examples, but only part of the

---

[10] Even though our understanding of how the brain computes is still in its infancy, important breakthroughs in cortical (column) theory have been achieved in the last decade.
[11] For example, https://greenwaves-technologies.com/
[12] https://arxiv.org/abs/1706.03762

sequence, the remaining part being the "ground truth". The biggest models, such as GPT3, are based on this architecture. GPT3 was in the spotlights in May 2020 because of its potential use in many different applications (the context being given by the beginning sequence) such as generating new text, summarizing text, translating text, answering to questions and even generating code from specifications. This was even amplified by GPT4, and all those capabilities were made visible to the public in November 2022 with Chat-GPT, which triggered a maximum of hype and expectations. Even if transformers are mainly used for cloud applications today, this kind of architecture is rippling down in embedded systems. Small to medium size (1, 3, 7 to 13 G parameters models) can be executed on single board computers such as Jetson Orin Nano and even Raspberry PI. Quantization is a very important process to reduce the memory footprint of those models and 4-bit LLMs perform rather well. The new GPUs of NVIDIA support float8 in order to efficiently implement transformers. Supporting LLMs in a low-power and efficient way on edge devices is a new important challenge. However, most of these models are too big to run at the edge.

But we observe a trend in using several smaller specialized generative AI working together which could give comparable results than large monolithic LLMs. For example, the concept of Agentic AI, Mixture of Agents (MoA)[13], where a set of smaller Large Language Models (LLMs) works collaboratively. Smaller specialized LLMs can be trained and fine-tuned for specific tasks or domains. This specialization allows each model to become highly efficient and accurate in its particular area of expertise. By dividing the workload among these specialized agents[14], the system can leverage the strengths of each model, achieving a level of performance and accuracy that rivals or even surpasses that of a single, very large LLM. This targeted approach not only enhances the precision of the responses but also reduces the computational overhead associated with training and running a monolithic model. The MoA framework optimizes resource utilization by distributing tasks dynamically based on the specific strengths of each agent. This means that instead of overloading a single model with diverse and potentially conflicting requirements, the system can route tasks to the most suitable agent. Such an arrangement ensures that computational resources are used more efficiently, as each agent processes only the type of data it is best equipped to handle, and the others don't consume processing power, hence energy. Moreover, the MoA approach enhances scalability and maintainability and is more suited for edge devices, or a network of edge devices.

Training and updating a single enormous LLM is a complex and resource-intensive process, often requiring extensive computational power and time only available in the cloud or in large data centers. In contrast, updating smaller models can be more manageable and less resource-demanding. Additionally, smaller models can be incrementally improved or replaced without disrupting the entire system. This modularity allows for continuous enhancement and adaptation to new data and tasks, ensuring that the system remains current and effective over time. Furthermore, MoA systems inherently offer robustness and fault tolerance: with multiple smaller agents, the failure of one model does not cripple the entire system. This MoA (or similar) are very well suited for Edge AI by distributing the specialized models in different (parts of) systems, and having them interconnected (hence the requirement for edge hardware to have the connectivity stack integrated, with all security requirements).

[13] https://www.together.ai/blog/together-moa

[14] For example, using appraoches like RouteLLM, see https://lmsys.org/blog/2024-07-01-routellm/

Concerning the training of large foundation models, it is clear that this is only possible now in cloud AI due to the very large training dataset and computing power required to train foundation models. However, there are at least two other ways to adapt the foundation model to a particular (local) context: fine tuning and large token context. Fine-tuning involves taking a pre-trained foundation model and further training it on a specific, smaller dataset that reflects the target application's domain or context. This process adjusts a very small proportion of the model's weights (using LoRA or Adapters[15]) based on the new data, allowing it to specialize and perform more accurately in the given local context. This can be done on the edge with local data, for example in PC or smartphone which have sufficiently memory for storing the training data, RAM for running the fine-tuning process and of course processing power. This can be done when the system is idle, e.g. during the night for PCs or smartphone.

Model Size	Full Fine-tuning	LoRA	Q-LoRA
8B	60 GB	16 GB	6 GB
70B	300 GB	160 GB	48 GB
405B	3.25 TB	950 GB	250 GB

Note: These are estimated values and may vary based on specific implementation details and optimizations.

Figure 2.1.3: This table outlines the approximate memory requirements for training Llama 3.1 models using different techniques, from https://huggingface.co/blog/llama31

Leveraging a large token context refers to the model's ability to process and understand long sequences of text, providing it with a broader context window. This extended context allows the model to capture more nuanced dependencies and relationships within the data, improving its comprehension and relevance to specific local contexts. By accommodating a larger context, the model can maintain coherence and produce more accurate and contextually appropriate outputs over longer spans of text, making it particularly useful for applications that require in-depth understanding and continuity, such as detailed document summarization or extended conversational AI systems. Together, these methods enable foundation models to be effectively adapted and optimized for specialized applications, enhancing their utility and performance in specific scenarios and they can be usable (in a near future) at the Edge in Embedded AI systems. Enlarging the context mainly implies more memory, as shown in Figure 2.1.4.

---

[15] https://github.com/AGI-Edgerunners/LLM-Adapters

Model Size	1k tokens	16k tokens	128k tokens
8B	0.125 GB	1.95 GB	15.62 GB
70B	0.313 GB	4.88 GB	39.06 GB
405B	0.984 GB	15.38	123.05 GB

Figure 2.1.4: requirements for the KV cache memory of Llama 3.1 models (FP16), from https://huggingface.co/blog/llama31

Energy efficiency of AI training:

Training AI models can be very energy demanding. As an example, according to a recent study, the model training process for natural-language processing (NLP, that is, the sub-field of AI focused on teaching machines to handle human language) could end emitting as much carbon as five cars in their lifetimes[16], [17]. However, if the inference of that trained model is executed billions of times (e.g. by billion users' smartphones), its carbon footprint could even offset the training one. Another analysis[18], published by the OpenAI association, unveils a dangerous trend: "since 2012, the amount of compute used in the largest AI training runs has been increasing exponentially with a 3.5 month-doubling time (by comparison, Moore's law had a 2-years doubling period)". These studies reveal that the need for computing power (and associated power consumption) for training AI models is dramatically widening. Consequently, the AI training processes need to turn greener and more energy efficient.

[16] https://www.technologyreview.com/2019/06/06/239031/training-a-single-ai-model-can-emit-as-much-carbon-as-five-cars-in-their-lifetimes/

[17] However, technology and algorithms improved a lot over time, and training a similar model (Bloom) 3 years later required 21 times less CO_2 emissions.

[18] https://openai.com/index/ai-and-compute/

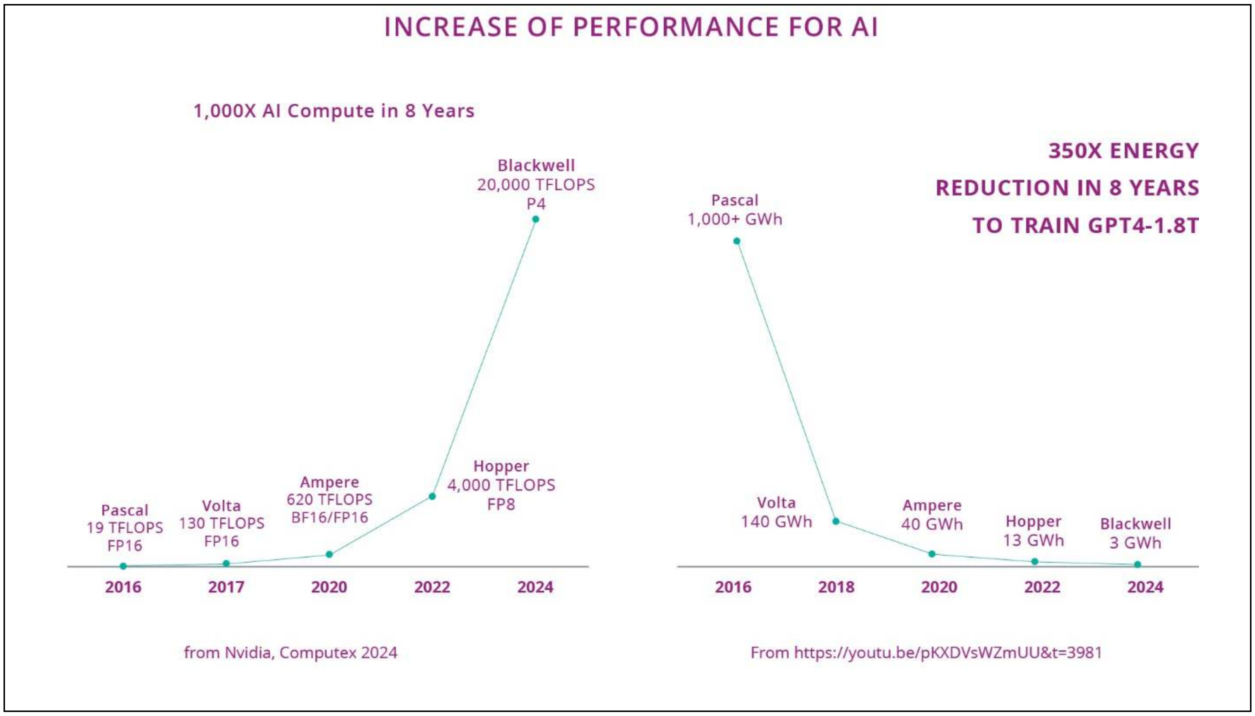


Figure 2.1.5: Increase of performance for AI

In fact, we can say that it is the available performance in operation per Watt of AI accelerators (GPUs) that drive the evolution of AI based on neural networks, including generative AI: the Figure 2.1.5 shows that GPT-4, launched on March 14, 2023, was only possible because the cost of the energy to train it in 2022 was acceptable and it would not have been possible to train it before. The Figure 2.1.5 also shows that GPU performances increased by 3 decades in 8 years, both thanks to new architecture, reduced data type (from FP16 to FP4) and smaller technology nodes. In this same period of time, the energy reduction was 350x.

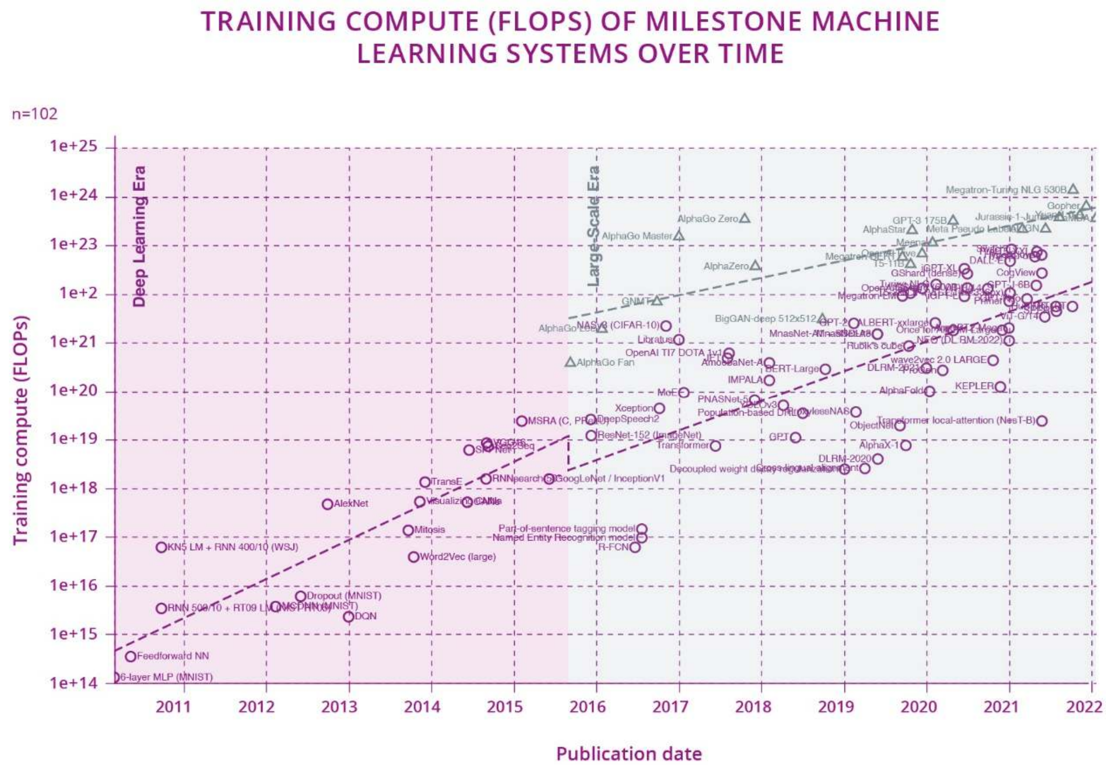


Figure 2.1.6:Evolution of the size of the most advanced deep learning networks (from https://arxiv.org/abs/2202.05924)

For a given use-case, the search for the optimal solution should meet multi-objective trade-offs among accuracy of the trained model, its latency, safety, security, and the overall energy cost of the associated solution. The latter means not only the energy consumed during the inference phase but also considering the frequency of use of the inference model and the energy needed to train it.

In addition, novel learning paradigms such as transfer learning, federated learning, self-supervised learning, online/continual/incremental learning, local and context adaptation, etc., should be preferred not only to increase the effectiveness of the inference models but also as an attempt to decrease the energy cost of the learning scheme. Indeed, these schemes avoid retraining models from scratch all the time or reduce the number and size of the model parameters to transmit back and forth during the distributed training phase.

It is also important to be able to support LLMs at the edge, in a low-cost and low-energy way, to benefit from their features (natural language processing, multimodality, few shot learning, etc..). Applications using transformers (such as LLMs) can run with 4 bit – or less - for storing each parameter, allowing to reduce the amount of memory required to use them in inference mode.

Although significant efforts have been focused in the past to enable ANN-based inference on less powerful computing integrated circuits with lower memory size, today, a considerable challenge to overcome is that non-trivial Deep Learning (DL)-based inference requires significantly more than the 0.5-1 MB of SRAM, that is the typical memory size integrated on top of microcontroller devices. Several approaches and methodologies to artificially reduce the size of a DL model exist, such as quantizing the neural weights and biases or pruning the network layers. These approaches are fundamental also to reduce the power consumption of the inference devices, but clearly, they cannot represent the definitive solution of the future.

We witness great development activity of computing systems explicitly supporting novel AI-oriented use cases, spanning different implementations, from chips to modules and systems. Moreover, as depicted in the following figure, it covers large ranges of performance and power, from high-end servers to ultra-low power IoT devices.

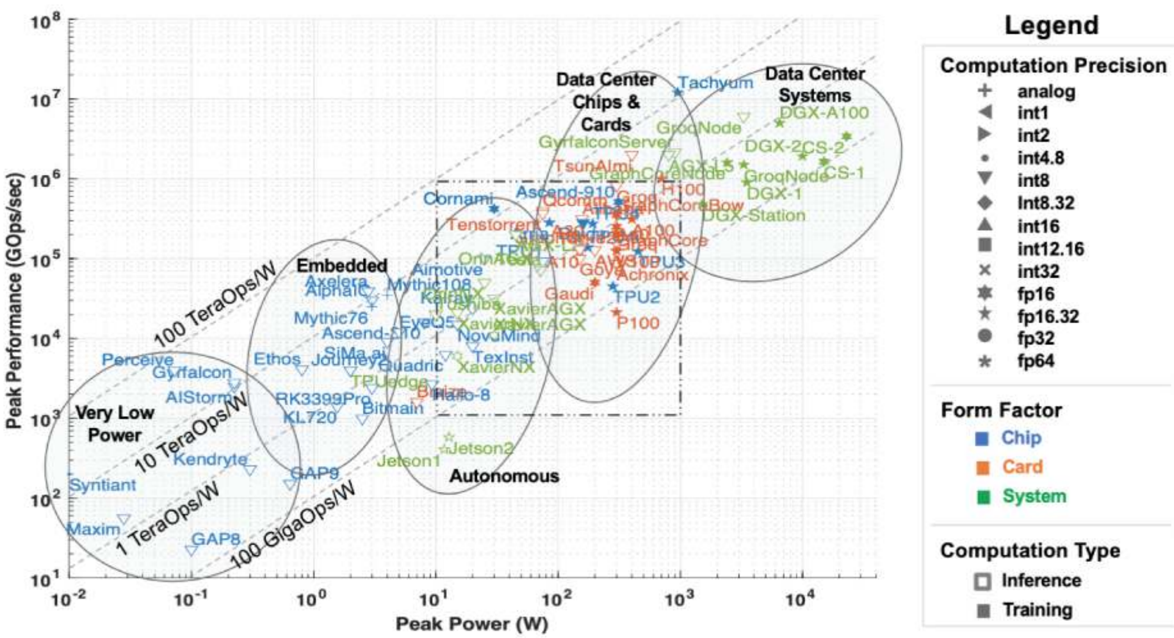


Figure 2.1.7: Landscape of AI chips according to their peak power consumption and peak performance[19].

To efficiently support new AI-related applications, for both, the server and the client on the edge side, new accelerators need to be developed. For example, DL does not usually need a 32/64/128-bit floating point for its learning phase, but rather variable precision including dedicated formats such as bfloats. However, a close connection between the compute and storage parts is required (Neural Networks are an ideal "compute in memory" approach). Storage also needs to be adapted to support AI requirements (specific data accesses, co-location compute and storage), memory hierarchy, local vs. cloud storage. This is particularly important for LLMs which (still) have a large number of parameters (few billions) to be efficient. Quantization into 4 to 2 bits, new memories and clever architectures are required for their efficient execution at the edge.

Similarly, at the edge side, accelerators for AI applications will particularly require real-time inference, in view to reduce the power consumption. For DL applications, arithmetic operations are simple (mainly multiply-accumulate) but they are done on data sets with a very large set of data and the data access is

[19] AI Accelerator Survey and Trends, Albert Reuther, Peter Michaleas, Michael Jones, Vijay Gadepally, Siddharth Samsi, Jeremy Kepner, October 2022 https://arxiv.org/abs/2210.04055

therefore challenging. In addition, clever data processing schemes are required to reuse data in the case of convolutional neural networks or in systems with shared weights. Computing and storage are deeply intertwined. And of course, all the accelerators should fit efficiently with more conventional systems.

Reducing the size of the neural networks and the precision of computation is key to allow complex deep neural networks to run on embedded devices. This can be achieved either by pruning the topology of the networks, and/or by reducing the number of bits storing values of weight and neuron values. These processes can be done during the learning phase, or just after a full precision learning phase, or can be done (with less performance) independently of the learning phase (example: post-training quantization). The pruning principle is to eliminate nodes that have a low contribution to the final result. Quantization consists either in decreasing the precision of the representation (from float32 to float16 or even float8, as supported by the NVIDIA GPUs mainly for transformer networks), or to change the representation from float to integers. For the inference phase, current techniques allow to use 8-bit representations with a minimal loss of performance, and sometimes to reduce the number of bits further, with an acceptable reduction of performance or small increase of the size of the network (LLMs still seem to have a good performance with a 4-bit quantization). Most major development environments (TensorFlow Lite[20], Aidge[21], etc.) support post-training quantization, and the Tiny ML community is actively using it. Supporting better tools and algorithms to reduce size and computational complexity of Deep Neural Networks is of paramount importance for allowing efficient AI applications to be executed at the edge.

Fixing and optimizing some parts of the processing (for example feature extraction for CNNs) leads to specialized architectures with very high-performance, as exemplified in the ANDANTE project.

[20] https://www.tensorflow.org/lite/performance/post_training_quantization
[21] https://projects.eclipse.org/projects/technology.aidge

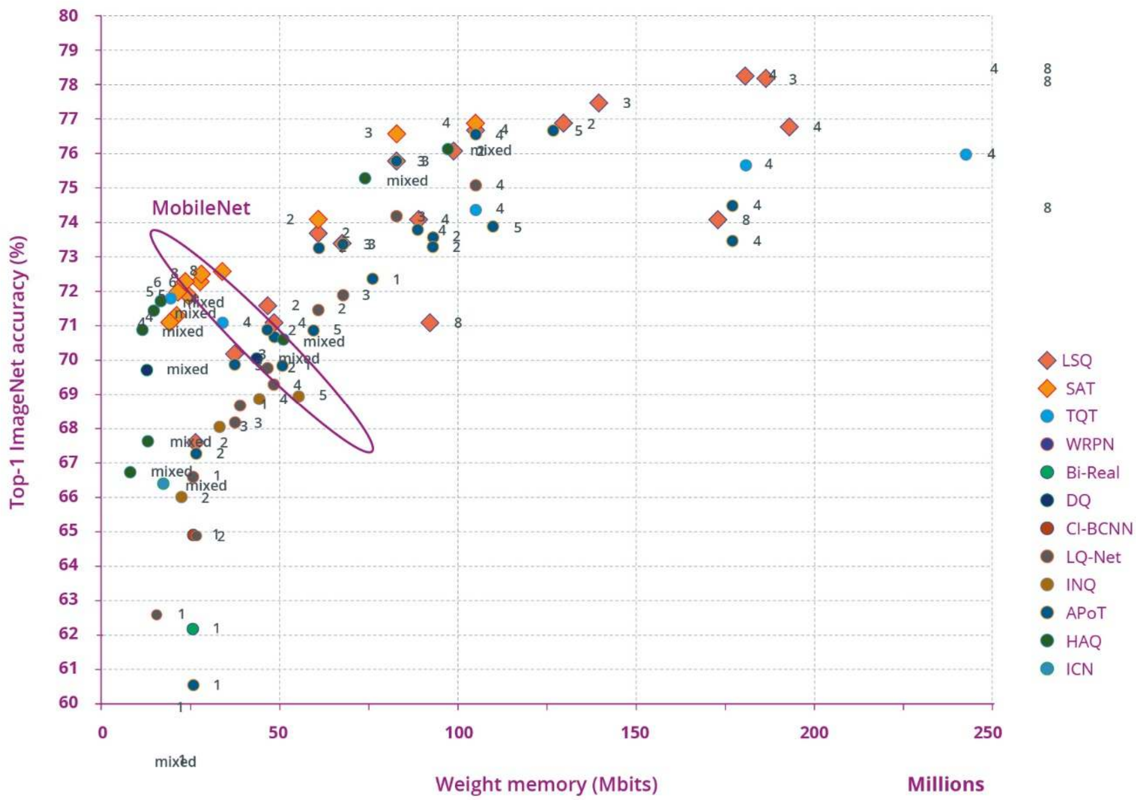


Figure 2.1.8: Results of various quantization methods versus Top-1 ImageNet accuracy

2.1.4.1 Impact of AI and embedded intelligence in sustainable development

Recently, the attention paid to the identification of sustainable computing solutions in modern digitalization processes has significantly increased. Climate changes and an initiative like the European Green Deal[22] are generating more sensitivity to sustainability topics, highlighting the need to always consider the technology impact on our planet, which has a delicate equilibrium with limited natural resources[23]. The computing approaches available today, as cloud computing, are in the list of the technologies that could potentially lead to unsustainable impacts. A study[24] has clearly confirmed the importance of edge computing for sustainability but, at the same time, highlighted the necessity of increasing the emphasis on sustainability, remarking that "research and development should include

[22] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en

[23] Nardi, B., Tomlinson, B., Patterson, D.J., Chen, J., Pargman, D., Raghavan, B., Penzenstadler, B.: Computing within limits. Commun. ACM. 61, 86–93 (2018)

[24] Hamm, Andrea & Willner, Alexander & Schieferdecker, Ina. (2020). Edge Computing: A Comprehensive Survey of Current Initiatives and a Roadmap for a Sustainable Edge Computing Development. 10.30844/wi_2020_g1-hamm.

sustainability concerns in their work routine" and that "sustainable developments generally receive too little attention within the framework of edge computing". The study identifies three sustainability dimensions (societal, ecological, and economical) and proposes a roadmap for sustainable edge computing development where the three dimensions are addressed in terms of security/privacy, real-time aspects, embedded intelligence and management capabilities.

AI and particularly embedded intelligence, with its ubiquity and its high integration level having the capability "to disappear" in the environment (ambient intelligence), is significantly influencing many aspects of our daily life, our society, the environment, the organizations in which we work, etc. AI is already impacting several heterogeneous and disparate sectors, such as companies' productivity[25], environmental areas like nature resources and biodiversity preservation[26], society in terms of gender discrimination and inclusion[27] [28], smarter transportation systems[29], etc. just to mention a few examples. The adoption of AI in these sectors is expected to generate both positive and negative effects on the sustainability of AI itself, on the solutions based on AI and on their users[30] [31]. It is difficult to extensively assess these effects and there is not, to date, a comprehensive analysis of their impact on sustainability. A study[32] has tried to fill this gap, analyzing AI from the perspective of 17 Sustainable Development Goals (SDGs) and 169 targets internationally agreed in the 2030 Agenda for Sustainable Development[33]. From the study it emerges that AI can enable the accomplishment of 134 targets, but it may also inhibit 59 targets in the areas of society, education, health care, green energy production, sustainable cities, and communities.

From a technological perspective AI sustainability depends, at first instance, on the availability of new hardware and software technologies. From the application perspective, automotive, computing and healthcare are propelling the large demand of AI semiconductor components and, depending on the application domains, of components for embedded intelligence and edge AI. This is well illustrated by car factories being on hold because of the shortage of electronic components. Research and industry organizations are trying to provide new technologies that lead to sustainable solutions redefining

---

[25] Acemoglu, D. & Restrepo, P. Artificial Intelligence, Automation, and Work. NBER Working Paper No. 24196 (National Bureau of Economic Research, 2018). https://futurium.ec.europa.eu/en/connect-university/events/next-computing-paradigm-hipeac-2024

[26] Norouzzadeh, M. S. et al. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. Proc. Natl Acad. Sci. USA 115, E5716–E5725 (2018).

[27] Bolukbasi, T., Chang, K.-W., Zou, J., Saligrama, V. & Kalai, A. Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. Adv. Neural Inf. Process. Syst. 29, 4349–4357 (2016).

[28] Tegmark, M. Life 3.0: Being Human in the Age of Artificial Intelligence (Random House Audio Publishing Group, 2017)

[29] Adeli, H. & Jiang, X. Intelligent Infrastructure: Neural Networks, Wavelets, and Chaos Theory for Intelligent Transportation Systems and Smart Structures (CRC Press, 2008).

[30] Jean, N. et al. Combining satellite imagery and machine learning to predict poverty. Science (80-.) 353, 790–794 (2016)

[31] Courtland, R. Bias detectives: the researchers striving to make algorithms fair. Nature 558, 357–360 (2018).

[32] Vinuesa, R., Azizpour, H., Leite, I. et al. The role of artificial intelligence in achieving the Sustainable Development Goals. Nat Commun 11, 233 (2020).

[33] UN General Assembly (UNGA). A/RES/70/1Transforming our world: the 2030 Agenda for Sustainable Development. Resolut 25, 1–35 (2015).

traditional processor architectures and memory structure. We already saw that computing near, or in-memory, can lead to parallel and high-efficient processing to ensure sustainability.

The second important component of AI that impacts sustainability concerns software and involves the engineering tools adopted to design and develop AI algorithms, frameworks, and applications. The majority of AI software and engineering tools adopt an open-source approach to ensure performance, lower development costs, time-to-market, more innovative solutions, higher design quality and software engineering sustainability. However, the entire European community should contribute and share the engineering efforts at reducing costs, improving the quality and variety of the results, increasing the security and robustness of the designs, supporting certification, etc.

The report on "Recommendations and roadmap for European sovereignty on open-source hardware, software and RISC-V Technologies"[34] discusses these aspects in more details.

Eventually, open-source initiatives (being so numerous, heterogeneous, and adopting different technologies) provide a rich set of potential solutions, allowing to select the most sustainable one depending on the vertical application. At the same time, open source is a strong attractor for application developers as it gathers their efforts around the same kind of solutions for given use cases, democratizes those solutions and speeds up their development. However, some initiatives should be developed, at European level, to create a common framework to easily develop different types of AI architectures (CNN, ANN, SNN, LLM, etc.). This initiative should follow the examples of GAMAM (Google, Amazon, Meta, Apple, Microsoft). GAMAM have greatly understood its value and elaborated business models in line with open source, representing a sustainable development approach to support their frameworks[35]. It should be noted that open-source hardware should not only cover the processors and accelerators, but also all the required infrastructure IPs to create embedded architectures. It should be ensured that all IPs are interoperable and well documented, are delivered with a verification suite, and remain maintained constantly to keep up with errata from the field and to incorporate newer requirements. The availability of automated SoC composition solutions, allowing to build embedded architectures design from IP libraries in a turnkey fashion, is also a desired feature to quickly transform innovation into PoC (Proof of Concept) and to bring productivity gains and shorter time-to-market for industrial projects.

The extended GAMAM and the BATX[36] also have large in-house databases required for the training and the computing facilities. In addition, almost all of them are developing their chips for DL (e.g. Google with its line of TPUs) or made announcements that they will. The US and Chinese governments have also started initiatives in this field to ensure that they will remain prominent players in the field, and it is a domain of competition.

Sustainability through open technologies extends also to open data, rules engines[37] and libraries. The publication of open data and datasets is facilitating the work of researchers and developers for ML and DL, with the existence of numerous images, audio and text databases that are used to train the models

---

[34] https://digital-strategy.ec.europa.eu/en/library/recommendations-and-roadmap-european-sovereignty-open-source-hardware-software-and-risc-v

[35] DL networks with Tensorflow at Google, PyTorch / Caffe at Facebook, CNTK at Microsoft, Watson at IBM, DSSTNE at Amazon

[36] BATX is an acronym standing for Baidu, Alibaba, Tencent, and Xiaomi, the four biggest tech firms in China

[37] Clips, Drools distributed by red Hat, DTRules by Java, Gandalf on PH

and become benchmarks[38]. Reusable open-source libraries[39] allow to solve recurrent development problems, hiding the technical details and simplifying the access to AI technologies for developers and SMEs, maintaining high-quality results, reducing time to market and costs.

The Tiny ML community (https://www.tinyml.org/) is bringing Deep Learning to microcontrollers with limited resources and at ultra-low energy budget. The MLPerf allows to benchmark devices on similar applications (https://github.com/mlcommons/tiny), because it is nearly impossible to compare performances on figures given by chips providers. Other Open-source initiatives are gearing towards helping the adaptation of AI algorithms to embedded systems, such as the Eclipse Aidge[40] , supported by the EU founded project Neurokit2E[41].

From the application perspective, impact of AI and embedded intelligence on sustainable development could be important. By integrating advanced technologies into systems and processes, AI and embedded intelligence enhance efficiency, reduce resource consumption, and promote sustainable practices, ultimately contributing to the achievement of the United Nations' Sustainable Development Goals (SDGs).

AI-powered systems could be used to optimizing energy consumption and improving the efficiency of renewable energy sources. For example, smart grids utilize AI to balance supply and demand dynamically, integrating renewable energy sources like solar and wind more effectively. Predictive analytics enable better forecasting of energy production and consumption patterns, reducing waste and enhancing the reliability of renewable energy. Additionally, AI-driven energy management systems in buildings and industrial facilities can significantly reduce energy usage by optimizing heating, cooling, and lighting based on real-time data.

Finally, embedded intelligence in consumer products promotes sustainability by enabling smarter usage and longer lifespans. For example, smart appliances can optimize their operations to reduce energy consumption, while predictive maintenance in industrial machinery can prevent breakdowns and extend equipment life, reducing the need for new resources and lowering overall environmental impact.

2.1.4.2 Market perspectives

Several market studies, although they don't give the same values, show the huge market perspectives for AI use in the next years.

According to ABI Research, it is expected that 1.2 billion devices capable of on-device AI inference have been shipped in 2023, with 70% of them coming from mobile devices and wearables. The market size for ASICs responsible for edge inference is expected to reach US$4.3 billion by 2024 including embedded architectures with integrated AI chipset, discrete ASICs, and hardware accelerators.

---

[38] A few examples are ImageNet (14 million images in open data), MNIST or WordNet (English linguistic basis)
[39] Nvidia Rapids, Amazon Comprehend, Google NLU Libraries
[40] https://projects.eclipse.org/projects/technology.aidge
[41] https://www.neurokit2e.eu/

The market for semiconductors powering inference systems will likely remain fragmented because potential use cases (e.g. facial recognition, robotics, factory automation, autonomous driving, and surveillance) will require tailored solutions. In comparison, training systems will be primarily based on traditional CPUs, GPUs, FPGAs infrastructures and ASICs.

According to McKinsey, it is expected that by 2025 AI-related semiconductors could account for almost 20 percent of all demand, which would translate into about $65 billion in revenue with opportunities emerging at both data centers and the edge.

According to a recent study, the global AI chip market was estimated to USD 9.29 billion in 2019 and it is expected to grow to USD 253.30 billion by 2030, with a CAGR of 35.0% from 2020-2030.

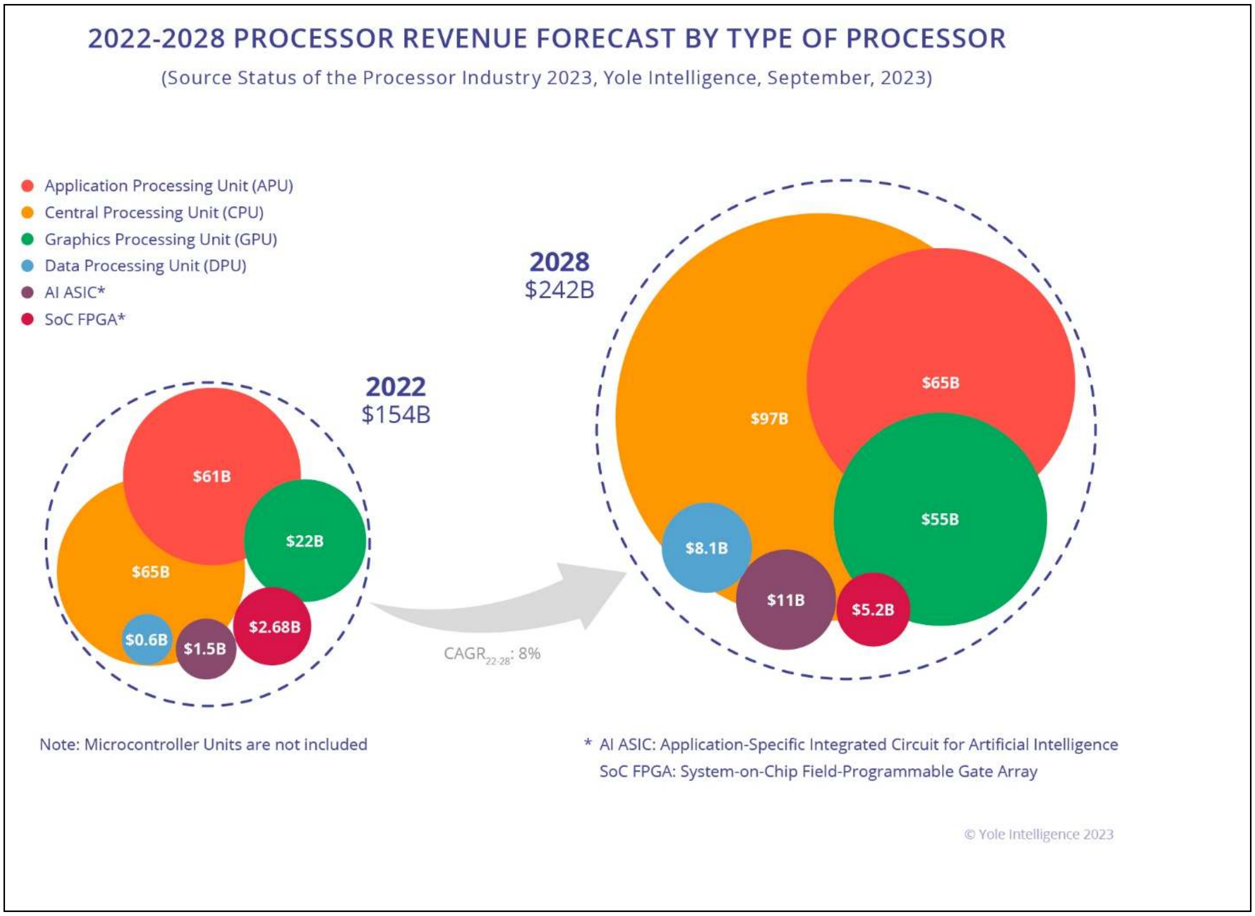


Figure 2.1.9: Market of high-end chips from https://www.yolegroup.com/product/report/status-of-the-processor-industry-2023/

There is a large increase in value of the market in the field of chip for IA, mainly for server/datacenters, but the market for Edge IA is also forecasted to have a large increase in the next years. The exact figures vary according to the market forecast company, but there is a common agreement for the large growth.

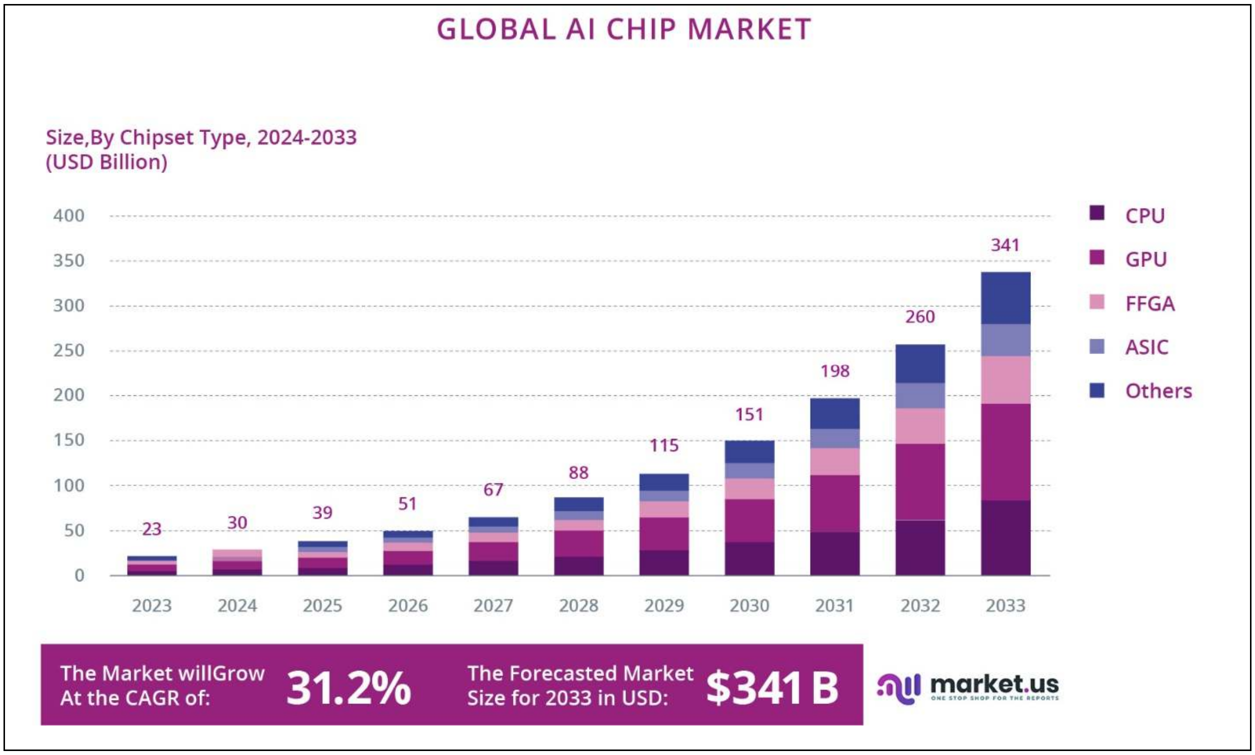


Figure 2.1.10: Market forecast of AI chips, from https://market.us/report/ai-chip-market/

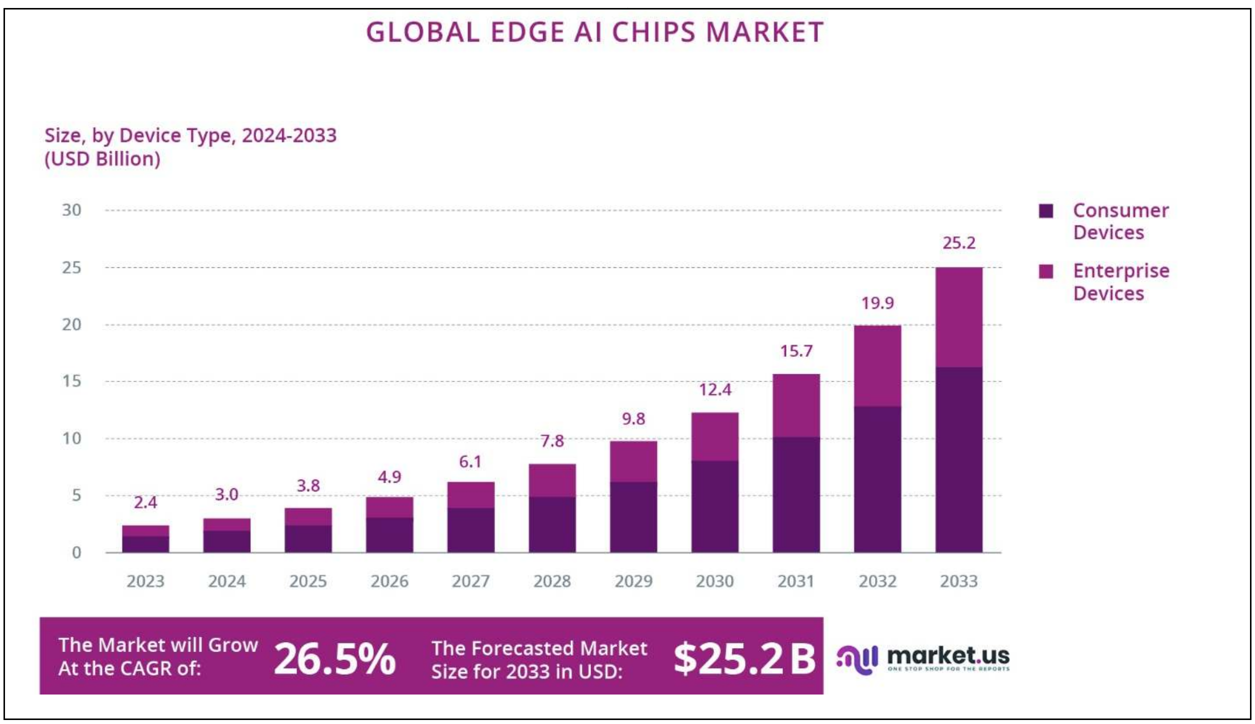


Figure 2.1.11: Edge AI market growth 2023-2033, from https://market.us/report/edge-ai-chips-market/

However, there is now an increasing trend towards AI chips that can execute generative AI models at the edge, and this trend is clearly increasing in 2024. These models can generate new content, predict complex sequences, and enhance decision-making processes in more sophisticated ways. Initially, the integration of these capabilities is seen in the PC and smartphone markets (Copilot+PC, with currently Snapdragon X series of chips, smartphones with Snapdragon 8 Gen 3 able to support generative AI models with up to 10 billion parameters[42], or the Mediatek Dimensity 9300[43]), where there is a growing demand for advanced AI applications that require both high performance and low latency. For instance, new smartphones and PCs are being equipped with AI chips capable of running generative models locally, enabling features such as real-time language translation, advanced virtual assistants, and enhanced creative tools like automated photo and video or text editing[44].

European semiconductor companies are not in this market of chipsets for smartphone or PCs, but **it is expected that, with the progress in the size reduction and specialization of large Foundation Models, generative AI at the edge will be used in the near (?) future in more and mode domains, including the ones primarily targeted by European chip manufacturers.** We also observe emerging attempts (not very successful yet) to develop new devices and use cases in the domain of edge devices (often connected to the cloud) such as Rabbit R1, Ai-pin, etc.

Another recent trend is that the **RISC-V instruction set based chips are also developing rapidly**, especially in China or Taiwan. For example, the Chinese company Sophgo delivers a cheap microcontroller with 2 RISC-V (one running at 1GHz, the other at 700 MHz) and a 1 TOPS (8bit) NPU, allowing to build systems running both Linux and a RTOS for less than 5$. On the cheap side, the CH32V003 is a 32-bit RISC-V microcontroller running at 48 MHz with 16kB of flash and 2 kB or SRAM which is sold at retail price (not for large quantities) below 10 cents. On the other side of the spectrum, various Chinese companies are announcing high performance RISC-V chips for servers, for example Alibaba with its C930[45]. US companies are also present in this market (SiFive, Esperanto, Meta – MTIA chip, etc). European companies are also involved in developing or using RISC-V cores (Codasip, Greenwaves, Bosch, Infineon, NXP, Nordic semiconductors, …). Research firm Semico projects a staggering 73.6 percent annual growth in the number of chips incorporating RISC-V technology, with a forecast of 25 billion AI chips by 2027 generating a revenue of US $291 billion.

2.1.4.3 AI components vendors

In the next few years, the hardware is serving as a differentiator in AI, and AI-related components will constitute a significant portion of future demand for different applications.

Qualcomm has launched a generation of Snapdragon processors, with NPU, allowing to run models up to 10 billion parameters in smartphones (such as Samsung Galaxy S24) or in PCs (Copilot+PCs), MediaTek

[42] https://www.qualcomm.com/products/mobile/snapdragon/smartphones/snapdragon-8-series-mobile-platforms/snapdragon-8-gen-3-mobile-platform#Overview
[43] https://www.mediatek.com/blog/whats-new-in-the-mediatek-dimensity-9300
[44] https://www.apple.com/apple-intelligence/
[45] https://www.tomshardware.com/pc-components/cpus/alibaba-claims-it-will-launch-a-server-grade-risc-v-processor-this-year

also has chips allowing to run generative AI on smartphones and Google, with its Tensor G4 (for the pixel 9) also introduces chips able to run small models in consumer grade (embedded) devices. Apple also introduces generative AI in their devices, starting from the iPhone 15 pro and device equipped with the M series of processors. Of course, the requirement is to have powerful NPUs (Copilot+PCs require NPU of 40 TOPs and 16 GB of RAM), but also a large amount of RAM able to host the model and the OS.

Huawei and MediaTek incorporate their embedded architectures into IoT gateways and home entertainment, and Xilinx finds its niche in machine vision through its Versal ACAP SoC. NVIDIA has advanced the developments based on the GPU architecture, NVIDIA Jetson AGX platform, a high performance SoC that features GPU, ARM-based CPU, DL accelerators and image signal processors. NXP and STMicroelectronics have begun adding AI HW accelerators and enablement SW to several of their microprocessors and microcontrollers.

ARM is developing core for machine learning applications and used in combination with the Ethos-U85[46] or Ethos-N78 AI accelerator. Both are designed for resource-constrained environments. The new ARM's cores are designed for customized extensions and for ultra-low power machine learning.

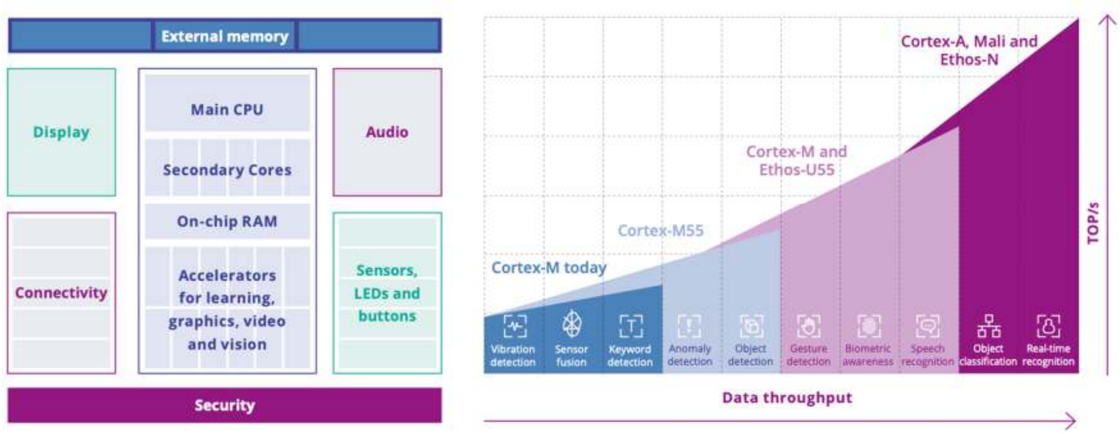


Figure 2.1.12: - Example of architecture of a modern SoC (from Paolo Azzoni, see also Chapter 1.3) / Arm's Cortex-M55 and Ethos-U55 Tandem. Provide processing power for gesture recognition, biometrics, and speech recognition applications (Source: Arm).

Companies like Google, Gyrfalcon, Mythic, NXP, STMicroelectronics and Syntiant are developing custom silicon for the edge. As an example, Google was releasing Edge TPU, a custom processor to run TensorFlow Lite models on edge devices. NVIDIA is releasing the Jetson Orin Nano range of products, allowing to perform up to 40 TOPS of sparce neural networks within a 15W power range[47].

[46] https://newsroom.arm.com/news/iot-reference-design-platform-2024
[47] https://developer.nvidia.com/blog/solving-entry-level-edge-ai-challenges-with-nvidia-jetson-orin-nano/

Open-source hardware, championed by RISC-V, will bring forth a new generation of open-source chipsets designed for specific ML and DL applications at the edge. French start-up GreenWaves is one of the European companies using RISC-V cores to target the ultra-low power machine learning space. Its devices, GAP8 and GAP9, use 8- and 9-core compute clusters, the custom extensions give its cores a 3.6x improvement in energy consumption compared to unmodified RISC-V cores.

The development of the neuromorphic architectures is accelerated as the global neuromorphic AI semiconductor market size is expected to grow.

The major European semiconductor companies are already active and competitive in the domain of AI at the edge:

- Infineon is well positioned to fully realize AI's potential in different tech domains. By adding AI to its sensors, e.g. utilizing its PSOC microcontrollers and its Modus toolbox, Infineon opens the doors to a range of application fields in edge computing and IoT. First, Predictive Maintenance: Infineon's sensor-based condition monitoring makes IoT work. The solutions detect anomalies in heating, ventilation, and air conditioning (HVAC) equipment as well as motors, fans, drives, compressors, and refrigeration. They help to reduce breakdowns, maintenance costs and extend the lifetime of technical equipment. Second, Smart Homes and Buildings: Infineon's solutions make buildings smart on all levels with AI-enabled technologies, e.g. building's domains such as HVAC, lighting or access control become smarter with presence detection, air quality monitoring, default detection and many other use cases. Infineon's portfolio of sensors, microcontrollers, actuators, and connectivity solutions enables buildings to collect meaningful data, create insights and take better decisions to optimize its operations according to its occupants' needs. Third, Health and Wearables: the next generation health and wellness technology is enabled to utilize sophisticated AI at the edge and is empowered with sensor, compute, security, connectivity, and power management solutions, forming the basis for health-monitoring algorithms in lifestyle and medical wearable devices supplying highest precision sensing of altitude, location, vital signs, and sound while also enabling lowest power consumption. Fourth, Automotive: AI is enabled for innovative areas such as eMobility, automated driving and vehicle motion. The latest microcontroller generation AURIX™ TC4x with its Parallel Processing Unit (PPU) provides affordable embedded AI and safety for the future connected, eco-friendly vehicle.

- NXP, a semiconductor manufacturer with strong European roots, has begun adding Al HW accelerators and enablement SW to several of their microprocessors and microcontrollers targeting the automotive, consumer, health, and industrial market. For automotive applications, embedded AI systems process data coming from the onboard cameras and other sensors to detect and track traffic signs, road users and other important cues. In the consumer space the rising demand for voice interfaces led to ultra-efficient implementations of keyword spotters, whereas in the health sector AI is used to efficiently process data in hearing aids and smartwatches. The industrial market calls for efficient AI implementations for visual inspection of goods, early onset fault detection in moving machinery and a wide range of customer specific applications. These diverse requirements are met by pairing custom accelerators, multipurpose and efficient CPUs with flexible SW tooling to support engineers implementing their system solution.

- STMicroelectronics integrated edge AI as one of the main pillars of its product strategy plan. By combining AI-ready features in its hardware products to a comprehensive ecosystem of software and tools, ST ambitions to overcome the uphill challenge of AI: opening technology access to all and for a broad range of applications. STMicroelectronics delivers a stack of Hardware and Software specifically designed to enable Neural Networks and Machine Learning inferences in an extremely low-energy environment. Most recently, STMicroelectronics has started to bring MCUs and MPUs with Neural accelerators to market, able to handle workloads that were not conceivable in an edge device a few years ago. These devices rely on NPUs (Neural Processing Units) that execute Neural Network tasks up to 100 times faster than in traditional high-end MCUs. The STM32MP2 MPU has been made available in the second quarter of 2024, it will be followed by a full family of STM32 MCUs leveraging STMicroelectronics home-grown NPU technology, Neural Art. The first MCU leveraging this technology is already in the hands of tens of clients for workloads such as visual events recognition, people and objects recognition, all executed at the edge in the device. On the software side, to adapt algorithms to small CPU and memory footprints, STMicroelectronics has delivered the ST Edge AI Suite, a toolset specifically designed to address all the needs of embedded developers from ideation with a rich model zoo, to datalogging to optimization for the embedded world of pre-created Neural Networks to creation of predictive maintenance algorithms. Two tools are particularly interesting: 1) STM32Cube.ai, an optimizer tool which enables a drastic reduction of the power consumption while maintaining the accuracy of the prediction. 2) NanoEdge AI studio, an Auto-ML software for edge-AI, that automatically finds and configures the best AI library for STM32 microcontroller or smart MEMS that contain ST's embedded Intelligent Sensor Processing Unit (ISPU). NanoEdge AI algorithms are widely used in projects such as arc-fault or technical equipment failure detection and extends the lifetime of industrial machines.

2.1.5 Applications breakthroughs

2.1.5.1 Example of AI application ranging from deep-edge to cloud

One example of an application that can leverage all three categories of AI hardware simultaneously in a true example of "continuum of computing" is an advanced autonomous driving system in smart cities.

In the cloud, massive amounts of data from numerous autonomous vehicles, traffic cameras, and other city infrastructure are aggregated and used to train large-scale AI models. Data centers equipped with GPUs handle the complex tasks of training models with large language models (LLMs). These models analyze patterns, improve decision-making algorithms, and enhance predictive capabilities. They also ensure that the autonomous driving algorithms continuously learn and adapt from real-world data. This centralized processing allows for the development of highly sophisticated models that can then be deployed to vehicles for local inferencing, ensuring they are always operating with the latest intelligence.

In the vehicles themselves, high-performance NPUs handle real-time inferencing tasks based on smaller models. These processors power the Advanced Driver Assistance Systems (ADAS) within each vehicle, enabling real-time perception, decision-making, and control. This includes tasks such as object detection, lane keeping, obstacle avoidance, and adaptive cruise control. This in-vehicle processing ensures that the car can react instantly to its environment without the latency that would be introduced by relying solely on cloud-based computations. They also ensure safety by working without a permanent connection to the cloud, which cannot be always guaranteed.

Within the vehicle, low-power AI hardware is used for continuous monitoring and processing of data from various sensors. These sensors include in-cabin cameras and audio systems, external cameras, LIDAR, radar, and ultrasonic sensors. The low-power AI systems process data for applications like driver behavior monitoring, passenger safety features, and environmental awareness (e.g., detecting nearby pedestrians or cyclists). By using low-power AI hardware, the vehicle can efficiently manage sensor data without draining the battery, which is crucial for electric vehicles. This also ensures that continuous monitoring and safety features remain active without interruption.

The integration of these three AI hardware categories creates a robust autonomous driving ecosystem. Cloud AI ensures that the overall traffic management in the city/countryside is well managed, and that the "intelligence" of the autonomous driving system is continuously improving through extensive data analysis and model training. High-Performance Embedded AI within the vehicle handles immediate, critical decision-making and control functions, allowing the car to operate safely and effectively in real time. Low-Power Embedded AI maintains constant environmental and internal monitoring, ensuring that all sensor data is processed efficiently, that the vehicle can respond to changing conditions without excessive power consumption, and that failures are prevented by continuous monitoring of the various functionalities of the vehicle.

It seems clear that advanced AI capabilities can enhance autonomous driving systems, in-car entertainment, and real-time vehicle diagnostics. Generative AI can be used to predict traffic patterns, generate detailed maps, simulate various driving scenarios to improve safety and efficiency, and can improve the interface between the car and the user.

2.1.5.2 Other application domains

What we observe in the automotive domain can also be applied to other domains of edge computing and embedded Artificial Intelligence, such as systems for factories, for homes, etc. As we have seen, **there is more and more convergence between edge computing and embedded (generative) AI, but still a lot of edge applications will be without AI (for now?), so AI support/accelerators are required only in few systems.**

Another significant market is the healthcare sector. AI chips capable of executing advanced AI models can be used in medical devices for helping diagnostics (with continuous monitoring, and forecasting potential problems), personalized treatment planning, and real-time monitoring of patient health. These applications require the ability to process and analyze data locally, providing immediate and accurate insights without relying on cloud-based solutions that may introduce latency or privacy concerns.

Important examples are its contribution in the recovery from the Covid-19 pandemic as well as its potential to ensure the required resilience in future crises[48].

Industrial automation and robotics also stand to benefit from this technology. Generative AI can enable more advanced predictive maintenance, optimize manufacturing processes, and allow robots to learn and adapt to new tasks more efficiently. By embedding AI chips with generative capabilities, industries can improve productivity, reduce downtime, and enhance operational flexibility. We see the first impressive results of using generative IA in robotics, not only for interfacing in natural ways with human, but also for scene analysis and action planning.

Internet of Things (IoT), smart home devices, wearable technology, and even agriculture technology can leverage generative AI for a variety of innovative applications. These include creating more intelligent home automation systems, developing wearables that provide more personalized health and fitness insights, and implementing smart farming techniques that can predict crop yields and optimize resource use.

While AI-accelerated chips have traditionally focused on supporting CNNs for image, audio, and signal analysis, the rise of generative AI capabilities is driving a new wave of AI chips designed for edge computing. This evolution is starting in the PC and smartphone markets but will spread across a range of embedded markets, including automotive, healthcare, industrial automation, robotics, and IoT, significantly broadening the scope and impact of edge AI technologies. Therefore, **it is important that Europe continues to be recognized as a player in the market of embedded MCU and MPU by being prepared to introduce generative AI accelerators in their products, even cheap ones.**

2.1.6 Major challenges

The convergence between edge computing and embedded generative AI is becoming increasingly evident, but many edge applications currently operate without AI integration. Consequently, AI support and accelerators are only required in a limited number of systems at present. Most of the accelerators available today in edge computing systems are designed for perception applications, such as image and audio processing, which require significant advancements to achieve further efficiency gains. As technology progresses, it is expected that the size reduction and specialization of large foundation models will make generative AI at the edge feasible in a growing number of domains. This trend includes areas primarily targeted by European chip manufacturers, signaling a shift towards broader AI integration in various applications.

Europe must maintain its status as a key player in the embedded MCU (Microcontroller Unit) and MPU (Microprocessor Unit) markets by preparing to introduce generative AI accelerators in their products, even in more affordable devices. These systems should also be ready for software upgradability, and connectivity allowing modularity and collaborative performances. This readiness will ensure that European manufacturers can meet the future demand for edge AI capabilities. Additionally, the rapid development of RISC-V instruction set-based chips presents a significant opportunity for innovation and

---

[48] https://www.eenewsembedded.com/news/nxp-developing-neural-networks-identify-covid-19

competitiveness in the global semiconductor market. Embracing these advancements will be crucial for Europe to stay at the forefront of the evolving edge computing landscape, where the integration of generative AI and specialized accelerators will become increasingly commonplace.

Europe should also be ready for the smooth integration of the devices into the computing continuum. This will not only require extra features in the embedded devices (such as connectivity, covered in the connectivity chapter, security, covered in the security chapter) but also a global software stack certainly based on As a Service approach, the development of smart orchestrators and a global architecture view at system level. As such, the evolution towards the computing continuum is not a major challenge of this chapter, but rather a global challenge. We will only list here some aspects that need to be further developed:

Four Major Challenges have been identified for Europe to be an important player in the further development of computing systems, especially in the field of embedded AI architectures and edge computing:

1. **Increasing Energy Efficiency**
2. **Managing the Increasing Complexity of Systems**
3. **Supporting the Increasing Lifespan of Devices and Systems**
4. **Ensuring European Sustainability**

2.1.6.1 Major Challenge 1: Increasing the energy efficiency of computing systems and embedded intelligence

2.1.6.1.1 State of the art

Increasing energy efficiency is the results of progress at multiple levels:

Technology level:

At technology level (FinFet, FDSOI, silicon nanowires or nanosheets), technologies are pushing the limits to be ultra-low power. Technologies related to advanced integration and packaging have also recently emerged (2.5D, chiplets, active interposers, etc.) that open innovative design possibilities, particularly for what concerns tighter sensor-compute and memory-compute integration.

Device level:

At device level, several types of architectures are currently developed worldwide. The list is moving from the well-known CPU to some more and more dedicated accelerators integrated in Embedded architectures (GPU, NPU, DPU, TPU, XPU, etc.), providing accelerated data processing and management capabilities, which are implemented going from fully digital to mixed or fully analog solutions:

- Fully digital solutions have addressed the needs of emerging application loads such as AI/DL workloads using a combination of parallel computing (e.g. SMP[49] and GPU) and accelerated hardware primitives (such as systolic arrays), often combined in heterogeneous Embedded architectures. Low-bit-precision (8-bit integer, 8- or 4-bits floating representations, …) computation as well as sparsity-aware acceleration have been shown as effective strategies to minimize the energy consumption per each elementary operation in regular AI/DL inference workloads; on the other hand, many challenges remain in terms of hardware capable of opportunistically exploiting the characteristics of more irregular mixed-precision networks. Some applications also require further development due to their need for more flexibility and precision in numerical representation (32- or 16-bit floating point), which puts a limit to the amount of hardware efficiency that can be achieved on the compute side.
- **Avoiding moving data:** this is crucial because the access energy of any off-chip memory is currently 10-100x more expensive than access to on-chip memory. Emerging non-volatile memory technologies such as MRAM, with asymmetric read/write energy cost, could provide a potential solution to relieve this issue, by means of their greater density at the same technology node. Near-Memory Computing (NMC) and In-Memory Computing (IMC) techniques move part of the computation near or inside memory, respectively, further offsetting this problem. While IMC in particular is extremely promising, careful optimization at the system level is required to really take advantage of the theoretical peak efficiency potential. Figure 2.1.13 shows that moving data requires an order of magnitude more energy than processing them (an operation on 64-bit data requires about 20pJ in 28nm technology, while getting the same 64-bit data from external DRAM takes 16nJ – without the energy of the external DRAM).

---

[49] Symmetric multiprocessing or shared-memory multiprocessing (SMP) involves a multiprocessor computer hardware and software architecture where two or more identical processors are connected to a single, shared main memory, from https://en.wikipedia.org/wiki/Symmetric_multiprocessing .

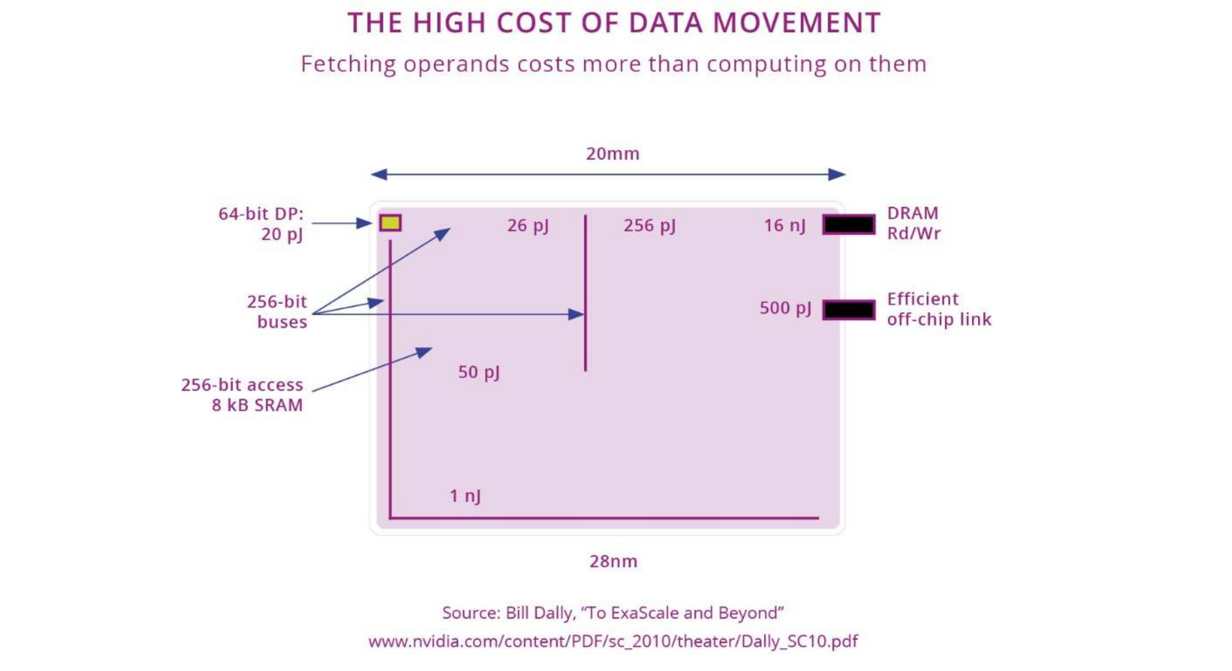


Figure 2.1.13:- Energy for compute and data movement. This explains the order of magnitude of the problem of data movement, and this problem is still relevant in all technology nodes.

This figure shows that in a modern system, large parts of the energy are dissipated in moving data from one place to another. For this reason, new architectures are required, such as computing in or near memory, neuromorphic architectures (where the physics of the NVM - PCM, CBRAM, MRAM, OXRAM, ReRAM, FeFET, etc. - technology can also be used to compute) and lower bit-count processing are of primary importance. Not only the memories itself, e.g. bitcells, are needed but the complementing libraries and IP for IMC or NMC as well.

		Flash reference	MRAM type	PCM type	ReRAM type	FeFET	Hf based FeRAM 1T1C
Performance	**Programming power**	<200pJ/bit - # 100pJ/bit (eSTM)	~20pJ/bit	~90pJ/bit	~100pJ/bit	<~20pJ/bit	<pJ/bit
	Reading access time	HV devices ~15ns	Core oxide device ~1ns	No HV devices ~5ns	No HV devices ~5ns	No HV devices ~5ns	Write after read (Destr. Read)
	Erasing	FN mechanis	bit-2-bit	bit-2-bit	bit-2-bit	bit-2-bit? Depend	bit-2-bit

	granularity	m Full page erasing	erasing Fine granularity	erasing Fine granularity	erasing Fine granularity	on archi	erasing
Reliability	**Endurance**	Mature technology	High capability 10^15 ?	500Kcy	10^5 trade Off with BER	10^5 Gate stress sensibility	10^11 #10^6 with write after read
	Retention	Mature technology	Main weakness Trade-off with Taa	150°C auto compliant	Demonstrated Trade Off with power	To be proven	To be proven
	Soldering reflow	Mature technology	High risk To be proven	pass	possible	To be proven	To be proven
Cost	**Extra masks**	Very high (>10)	Limited (3-5)	Limited (3-5)	Limited (3-5)	Low (1-3)	Low (1-3)
	Process flow	Complex	Complex	Simple	Simple	Simple	Simple
	New assets vs CMOS	Shared	New _ manufacturable	New _manufacturable	BE High-k material	FE High k material	BE High-k material

Figure 2.1.14: eNVM technologies, strengths and challenges (from Andante: CPS & IoT summer school, Budva, Montenegro, June 6th-10th, 2023

System level:

- Micro-edge computing near sensors (i.e. integrating processing inside or very close to the sensors or into local control) will allow embedded architectures to operate in the range of 10 mW (milliwatt) to 100 mW with an estimated energy efficiency in the order of 100s of GOPs/Watt up to a few TOPs/Watt in the next 5 years. This could be negligible compared to the energy consumption of the sensor (for example, a MEMS microphone can consume a few mW). On top, the device itself can go in standby or in sleep mode when not used, and the connectivity must not be permanent. Devices currently deployed on the edge rarely process data 24/7 like data centers: to minimize global energy, a key requirement for future edge Embedded architectures is to combine high performance “nominal” operating modes with lower-voltage high compute efficiency modes and, most importantly, with ultra-low-power sleep states, consuming well below

1 mW in fully state-retentive sleep, and less than 1-10 µW in deep sleep. The possibility to leave embedded architectures in an ultra-low power state for most of the time has a significant impact on the global energy consumed. The possibility to orchestrate and manage edge devices becomes fundamental from this perspective and should be supported by design. On the contrary, data servers are currently always on even if they are loaded only at 60% of their computing capability.

- **Unified (or shared) memory**, which allows CPUs and GPUs (or NPUs) to share a common memory pool, is gaining traction in first for high end systems such as Mx series of systems from Apple. This approach not only streamlines memory management but also provides flexibility by enabling the allocation of this unified memory pool to Neural Processing Units (NPUs) as needed. This is particularly advantageous for handling the diverse and dynamic workloads typical of AI applications and reduce the amount of energy required to move data between different parts of the system.
- **Maximum utilization and energy proportionality:** Energy efficiency in cloud servers rely on the principle that silicon is utilized at its maximum efficiency, ideally operating at 100% load. This maximizes the computational throughput per watt of power consumed, optimizing the energy efficiency of the server infrastructure. In cloud environments, high utilization rates are easier to achieve because workloads from multiple users can be aggregated and balanced across a vast number of servers. This dynamic allocation of resources ensures that the hardware is consistently operating near its capacity, minimizing idle periods. Considering processing at the edge, achieving a clear net benefit in terms of energy efficiency presents additional challenges. Unlike centralized cloud servers, edge devices often face fluctuations and unpredictable workloads. These devices must be capable of maintaining high efficiency across a wide range of utilization levels, not just at peak performance. To accomplish this, edge hardware needs to be designed with adaptive power management features that allow it to scale power usage dynamically in response to varying workloads. This means the **hardware should be as efficient at low workloads as it is at 100% utilization**, ensuring that energy is not wasted during idle or low-demand periods. Energy proportionality aims to scale computing power according to the computational demand at any given moment. This concept is crucial for modern computing environments, which must handle a diverse range of workloads, from the intensive processing needs of self-driving vehicles to the minimal computational requirements of IoT devices. The ability to dynamically adapt power usage ensures efficiency and effectiveness across this spectrum of applications. The **scalability of processing power is fundamental to achieving energy proportionality**. In high-demand scenarios, such as those faced by self-driving cars, the computational power must scale up to handle complex algorithms for navigation, sensor fusion, and real-time decision-making. Conversely, IoT devices often require minimal computational power, as they are designed to perform specific, limited functions with a few lines of code. This wide range of computational needs presents a significant challenge because the requirements and hardware for each application are vastly different.

Data level:

At data level, **memory hierarchies** will have to be designed considering the data reuse characteristics and access patterns of algorithms, which strongly impact load and store access rates and hence, the energy necessary to access each memory in the hierarchy and avoiding to duplicate data. For example (but not only), weights and activations in a Deep Neural Network have very different access patterns and can be deployed to entirely separate hierarchies exploiting different combinations of external Flash, DRAM, non-volatile on-chip memory (MRAM, FRAM, etc.) and SRAM.

The **right data type** should be also supported to perform correct (or as correct as acceptable) computation with the minimum number of bits and if possible, in integer (the floating-point representation as specified in IEEE 754 is very costly to implement, especially with the particular cases). AI based on transformers and generative AI could use a far simpler representation for the inference phase, such a bfloat16 (which could also be used for training) or even float 8 or 4 and even, in some cases, using less than 4 bits integer representation. It is obvious that storing and computing a weight coded with INT4 is about an order of magnitude more efficient than using a Float32 representation.

Tools level:

At tools level, **HW/SW co-design of system** and their associated algorithms are mandatory to minimize the data moves and optimally exploit hardware resources, particularly if accelerators are available, and thus optimize the power consumption. New AI-based HW/SW platforms with increased dependability, optimized for increased energy efficiency, low cost, compactness and providing balanced mechanisms between performance and interoperability should be further developed to support the integration into various applications across the industrial sectors of AI and other accelerators[50].

2.1.6.1.2 Key focus areas

Therefore, there are several key axes to further develop in order to increase energy efficiency of computing and AI systems:

Processing Data Locally and reducing data movements:

Increasing the energy efficiency of computing systems in general is a multifaceted challenge that requires innovations across various domains, including hardware, software, and data management. Indeed, power consumption should not only be seen at the level of the device, but at the levels composing the full technology stack. One effective strategy is **processing data locally where it is created**, which reduces the energy consumption associated with data transmission to centralized locations, and also reducing latency required in real-time applications, such as self-driving vehicles or other industrial commands. This localized data processing not only enhances energy efficiency but also helps in securing the data by keeping it within the local environment, thus ensuring privacy. Additionally, **the co-design of architectures, hardware, software, and topologies can significantly optimize system performance and energy usage**. This is closely linked to the concept of "local-first" software, as articulated by M. Kleppmann, A. Wiggins, P. Van Hardenberg, and M. McGranaghan in their 2019 paper, which presents a paradigm shift that emphasizes user autonomy and data ownership, even in the era of pervasive cloud computing. Local-first software is designed to ensure that users maintain control over their data by storing it primarily on their local devices, thus enabling full functionality even without an active internet connection. In the local-first paradigm, the software architecture is designed to prioritize local operations, using Conflict-free Replicated Data Types (CRDTs) and other synchronization mechanisms to manage concurrent data updates and ensure consistency across distributed systems. This approach aligns with the capabilities of modern edge hardware, which can handle complex data structures and algorithms locally,

---

[50] An example is the platform developed in the project Neurokit2E, see https://www.neurokit2e.eu/

reducing the reliance on centralized cloud infrastructures. As shown in the beginning of this chapter, this concept is clearly involved in Apple Intelligence and in HarmonyOS[51].

Co-Design of Algorithms, Hardware, Software:

The co-design of algorithms, hardware and software represents a holistic approach to optimizing system performance and energy usage. This integrated design process can help to identify and eliminate bottlenecks, streamline operations, and reduce overall energy consumption. It is also crucial to have a co-optimization of the software and hardware to explore more advanced trade-offs. At tools level, HW/SW co-design of system and their associated algorithms are mandatory to minimize the data moves and optimally exploit hardware resources, particularly if accelerators are available, and thus optimize the power consumption.

Efficient Management of storage resources:

In the landscape of **edge AI, the efficient management of memory resources is becoming increasingly critical to guarantee both a sustainable energy consumption profile and the required performances.** Local execution of smaller language models (SLMs) often encounters bottlenecks primarily at the memory level. This is particularly evident with the need for cheap and efficient memory solutions able not only to store but allowing to be used directly for computation, avoiding the transfer of the weight of the model from flash to standard DRAM – as done today – which is expensive. Research is only starting in this field[52]. Memory technology itself is advancing swiftly, with smartwatches now boasting up to 8GB of memory, showing the trend towards more substantial memory capacities in compact devices.

Moreover, the cost of memory remains a crucial factor for deploying generative AI at the edge. Efficient memory usage is imperative to keep costs manageable, especially as generative models require substantial memory for optimal performance**. Innovations in compressing weight** (quantization from FP16 to FP4 or even INT4 or lower), **pruning the network or advances in memory technology are therefore essential** to minimize RAM wastage and enhance the feasibility of deploying advanced AI capabilities on edge devices. In most NPUs, the weights have to be transferred from non-volatile memory to faster RAM, leading to energy waste and duplication of storage resources. This involves strategic allocation (between Flash, DRAM or other kind of memories) and utilization of memory to ensure that AI tasks are handled efficiently without unnecessary data copy (which consumes energy and requires more resources).

Concept of Unified Memory:

The concept of unified memory, where a common memory pool is shared between the CPU and GPU/NPU, offers significant advantages for energy efficiency. This approach allows for more flexible and dynamic allocation of memory resources, ensuring that the system can adapt to varying workloads. Unified memory can also help to reduce data duplication and transfer overheads, leading to more efficient processing.

---

[51] Some concepts are also developed in the open source Oniro project (https://oniroproject.org/), also based on OpenHarmony (https://gitee.com/openharmony).

[52] See https://arxiv.org/abs/2312.11514 for example

Innovations in Memory Technology:

Innovations in compressing weights, pruning neural networks, and advances in memory technology are essential for increasing energy efficiency. Techniques such as weight compression and network pruning reduce the amount of data that needs to be processed and stored, thereby lowering energy consumption. Advances in memory technology, such as the development of more efficient and high-capacity non-volatile memory solutions, that can be used directly for computing, further contribute to this goal.

Energy Proportionality:

Energy proportionality is a critical principle for hardware design, ensuring that devices are as efficient at low workloads as they are at 100% utilization. Addressing the scalability challenge involves **designing systems that can efficiently adjust their power usage without compromising performance**. This means developing hardware that is capable of scaling its power consumption dynamically and software that can efficiently distribute computational loads. The goal is to achieve a balance where each device, whether a powerful autonomous vehicle or a simple IoT sensor, operates at optimal energy efficiency.

Ultra-Low Standby Current:

Maintaining an ultra-low standby current is vital for devices that spend significant amounts of time in idle or low-power states, waiting to perform tasks triggered by specific events or intervals. This is particularly important for battery-operated and portable devices, which may include sensors, mobile devices, and IoT appliances, where energy efficiency is a key concern for these devices that spend significant amounts of time in standby mode. Achieving ultra-low standby current requires special power management techniques, such as power gating and dynamic voltage scaling, which can shut down non-essential components and reduce power supply to idle parts of the circuit without compromising the device's readiness to wake up and perform tasks promptly.

Leveraging Physical Phenomena for Computation:

One innovative approach to reducing energy consumption in computing involves **leveraging physical phenomena to perform computations**, effectively using the inherent properties of physical systems to solve complex problems. This method relies on the principle that certain physical processes can naturally execute the same types of equations and operations that we aim to solve computationally. This is the case for neuromorphic computing. This term covers various kind or realization, using sparce information coding method of "spikes", using RRAMs (resistive RAMs) to make essentially the weighted sum which is the basis of artificial neural network-based AI, including LLMs. Ohm's law can be used to make the (analog) product, and either Kirchhoff's law, or adding charges in a capacitor, for the summation. The value of RRAM can be also locally modified, allowing to implement some training algorithms very efficiently. However, despite their promising potential, neuromorphic computing technologies are still largely in the research phase. Scalability to large networks, dispersion of properties, interfacing with other systems, etc. still need to improve. **Extensive efforts are needed to refine these systems and validate their effectiveness** in real-world, product-ready solutions.

2.1.6.2 Major Challenge 2: Managing the increasing complexity of systems and embedded intelligence

2.1.6.2.1 State of the art

Managing the increasing complexity of systems is another crucial aspect. It is driven by the diversity and dependencies of its components as well as by the requested functions and performance, affecting concepts, architectures and operation at all levels of the system hierarchy. Therefore, the reference architectures for future AI-based systems need to provide modular and scalable solutions that support interoperability and interfaces among platforms that can collaborate, exchange information and share computing resources to allow the functional evolution of the silicon-born embedded systems, ready for the computing continuum evolution.

Still, it is observed that the strategical backbone technologies to realize such new architectures are not available. These strategical backbone technologies include smart and scalable electronic components and systems, the AI accelerator and control hardware and software, the security engines, and the connectivity technologies.

To achieve this, balanced mechanisms must be developed to ensure performance while maintaining interoperability between diverse systems. AI techniques play a vital role in complexity management by enabling self-optimizing, self-reconfiguring, and self-managing systems. These self-X capabilities allow systems to adapt dynamically to changing conditions, enhancing their reliability and resilience, but they are not available yet to the required extent.

The current SoCs are reaching a very high complexity by integrating a lot of functions and different accelerators on the same die. A decomposition of the SoCs into chiplets for more flexibility and scalability (e.g., mix and match) may help dealing with the complexity but for that an ecosystem which facilitates the design, and also provides plug and play capable, interoperable chiplets, is required.

A holistic end-to-end approach is necessary to manage the increasing complexity of systems, to remain competitive and to continuously innovate the European electronic components and systems ecosystem. This approach involves technologies described in other chapters such as AI for system conception described in methods and tools chapter. Besides theses, here is a short list of themes that need to be promoted in order to reduce global complexity of systems in general, and systems supporting AI techniques in particular:

- **Tools and techniques leveraging AI to help in the management of complexity**, e.g. tools to adapt LLMs to edge / embedded targets are required. For example, the development of AI-based HW/SW for multi-tasking and providing techniques to adapt the trained model to produce close or expected outputs when provided with a different but related set of data. The new solutions must provide dynamic transfer learning, by assuring the transfer of training instance, feature representation, parameters, and relational knowledge from the existing trained AI model to a new one that addresses the new target task. They should also ensure hyperparameter tuning for automated machine learning scenarios, etc. Furthermore, environments for **modeling how the various parts interact together should be considered since this topic is still a challenge;** having accurate digital twins of complex systems with the right level of abstraction for the various uses (conception, runtime optimization, etc.) is still an unsolved challenge.

- New design and architecture concepts including HW-SW codesign with corresponding HW/SW platforms for AI-born embedded systems should be promoted to facilitate trust by providing the dependable design techniques, that enable the end-to-end AI systems to be scalable, make correct decisions in a repetitive manner, provide mechanisms to be transparent, explainable, interpretable, and able to achieve interpretable results and embed features for AI model's and interfaces' interpretability. Linked to the previous point, infrastructure for the secure and safe execution of AI should be created.
- Enabling factors to **manage the complexity of the continuum of computing** are **the use and development of standardized APIs for hardware** and software tool chains, methods and interoperability across different system layers like sensors, gateways, on-premise servers and edge processing units. Additionally, interoperability concepts for AI edge-based platforms for data tagging, training, deployment, and analysis should be supported, allowing the development of distributed edge computing architecture with AI models running on distributed devices, servers, or gateways away from data centers or cloud servers. This implies scalable and hardware agnostic AI models capable of delivering comparable performance on different computing platforms, (e.g. Intel, AMD or ARM, Risc-V architectures) and seamless and secure integration at HW/SW embedded systems with the AI models integrated in the SW/HW and APIs to support configurable data integrated with enterprise authentication technologies through standards-based methods.
- **Deterministic behaviors, low latency and reliable communications** are also important for other vertical applications, such as connected cars, where edge computing and AI represent "the" enabling technology, independently from the sustainability aspects. The evolution of 5G is strongly dependent on edge computing and multi-access edge computing (MEC) developments.

In this chapter, five Key Focus Areas have been identified to tackle this complexity challenge:

- Complexity Management Utilizing AI
- Decomposition of Complex SoCs into Chiplets and Interposers
- Ensuring Programmability and Interoperability
- Combining Processing Devices to Work Together
- Modeling Interactions Among System Components

2.1.6.2.2 Key focus areas

Complexity management utilizing AI

In the domain of edge AI, the concept of a **Mixture of Agents** (MoA) or Agentic AI, where a set of smaller Large Language Models (LLMs) called agents works collaboratively, presents a promising approach to managing complexity more effectively than relying on very large "universal" foundation models. This strategy involves deploying multiple specialized models, each trained to handle specific tasks or domains, and orchestrating their cooperation to achieve comprehensive AI functionalities. By distributing the workload across several specialized agents, the system can maintain high levels of efficiency and accuracy, tailored to the particularities of different tasks. The MoA framework enhances scalability and flexibility, enabling more precise and context-aware responses by leveraging the strengths of each specialized model. This approach also allows for tailored optimization, enhancing the performance of AI applications across diverse domains. Additionally, it reduces the computational and memory overhead typically associated with very large models (only the relevant agents are active, not all of them), making it more feasible to implement advanced AI capabilities on edge devices with limited resources. Consequently, the MoA paradigm not only simplifies the management of AI complexity but also optimizes performance and resource utilization in edge computing environments, paving the way for more responsive and efficient AI applications.

Decomposition of Complex SoCs into Chiplets and Interposers

The decomposition of complex SoCs into chiplets and interposers is an emerging strategy to manage system complexity. This approach involves breaking down a monolithic SoC into smaller, modular components (chiplets) that are interconnected using interposers. This architecture allows for greater flexibility in design and manufacturing, enabling more efficient scaling and customization of systems, e.g. adaptation to domains/applications. Additionally, it facilitates better management of thermal and power characteristics, improving overall system performance and it may help to optimize cost. By adopting this decomposition approach featuring an ecosystem which includes interoperable chiplets, we can more effectively address the challenges posed by the increasing complexity of systems. However, for chiplets and interposer ecosystems to really emerge, an agreed standard for interconnect (such as UCIe) will be required, together with physical specifications allowing interoperability between providers.

Ensuring Programmability and Interoperability

To handle the increasing complexity of systems, it is essential to ensure programmability and interoperability for example using techniques and adherence to standards that originate from the web and cloud computing (but not necessarily involving the cloud). Technologies such as containers, orchestration, and WebAssembly (WASM) maintain high flexibility and adaptability and enable seamless integration and operation across different platforms managing the diversity and dynamic range of modern computing environments. These tools allow developers to manage and update systems more efficiently, ensuring that various components can communicate and function together effectively while developer's productivity can be helped by generative AI (chapter "Embedded Software" of this SRIA), e.g., for the exact syntax of the API for driving a particular function, or even generating templates to use a part of the hardware.

Combining Processing Devices to Work Together

The combination of various processing devices working together is another key focus area for managing system complexity. By combining CPUs, GPUs, NPUs, and other specialized processors systems can distribute tasks according to the strengths of each device. The combination may be either locally (integration at chip level, at the package level using for example chiplets or across distributed chips) or as separate components of a system, like in a car, or distributed on physically separated systems, each part having a special task to do, thus separating the concerns and helping the management of the complexity by splitting it explicitly. This collaborative approach ensures optimal performance and energy efficiency, as each processor type can handle specific aspects of the workload more effectively. Coordinating these devices requires sophisticated management and scheduling techniques, but it leads to more robust and scalable systems. Together with interoperability, the ability of devices to work together is the foundation of the computing continuum approach.

Modeling Interactions Among System Components

Modeling the interactions among various parts of a system remains a significant challenge in managing complexity. Understanding how different components interact, communicate, and influence each other is critical for optimizing system performance and reliability. Advanced modeling techniques and simulation tools are necessary to accurately represent these interactions and predict system behavior under different conditions. By improving our ability to model these interactions, we can better design and manage complex systems, ensuring they operate smoothly and efficiently.

2.1.6.3 *Major Challenge 3: Supporting the increasing lifespan of devices and systems and embedded intelligence*

2.1.6.3.1 State of the art

With the power of embedded AI, we can now unleash the self-X concept which didn't took off in the past. We can now design HW/SW techniques and architectures that can dynamically monitor their behavior, predict default (predictive maintenance), allowing devices to self-optimize, reconfigure and self-manage the resource demands (e.g. memory management, power consumption, AI model selection, etc…). This will have a direct effect in increasing the lifetime of devices and increasing their up-time.

Supporting the increasing lifespan of devices and systems can mainly be based on three main principles:

- Upgradability
- Modularity
- Reuse

By extension, this means that devices and systems have to increase their scalability and interoperability.

Upgradability

This requires hardware that can accommodate for example software upgrades, ensuring that **devices remain functional and up-to-date over extended periods**. This is the drive for approaches such as Software Defined Vehicle.

The first level of lifetime extension is clearly the upgrade to avoid replacing an object but instead improving its features and performance through either hardware or software update. This concept is not new as it has already been applied in several industrial domains for dozens of years.

Hardware requirements for edge processing devices that support software upgrades must address several critical factors to ensure that these devices remain functional and up-to-date over extended periods. The hardware must have sufficient computational power and memory to handle the evolving complexity of software applications. Storage capacity must be flexible and expandable, allowing for the installation of future updates and new applications without compromising performance. The devices must also incorporate robust connectivity options, such as advanced wireless protocols and multiple I/O ports, to ensure seamless integration with new peripherals and communication technologies that may emerge over time. This has an initial cost because the hardware is perhaps over-dimensioned for the exact initial requirements but is compensated for the user by its potential increased lifetime.

Secondly, hardware for edge processing devices must be compatible with various operating systems, RTOS and software ecosystems, providing the flexibility needed to adapt to different applications and use cases over time. This compatibility ensures that the devices can support a wide range of software upgrades and maintain interoperability with other systems. On the smartphone level, some companies (Google, Samsung), are guaranteeing that their device will be able to support software upgrade for 7 years. This is an important point that should also happen in the other domains of edge computing. But new disruptive innovations (like generative IA) can invalidate this long lifetime because of new drastic hardware requirements (mainly memory and requirement of a NPU for embedded generative IA).

Modularity

Enhancing interoperability and modularity allows different generations of devices to work together, extending their usability and reducing electronic waste. It would lead to systems which are more eco-conceived and with an augmented sustainability. The hardware should be **designed with modularity in mind**, enabling easy replacement or upgrading of individual components. This implies a long-lasting chip to chip communication protocol between units (like I_2C, SPI, PCI-e, USB, Ethernet, MIPI, AXI, etc.). **Modularity at the level of chiplets is not yet here** (adding/removing/upgrading chiplets). This modular approach not only extends the device's operational life by allowing for incremental hardware improvements but also reduces electronic waste and the overall cost of ownership.

Interoperability, modularity, scalability, virtualization, upgradability are well-known in embedded systems and are already widely applied. But they are brand new in AI and nearly non-existent in edge AI. On top, self-x (learning/training, configuration or reconfiguration, adaptation, etc.) are very promising but still under research or low level of development. Federative learning and prediction on the fly will certainly take a large place in future edge AI systems where many similar equipment collect data (Smartphone, electrical vehicles, etc.) and could be improved and refreshed continuously.

Reuse

The last aspect of increasing lifetime is to reuse a system in an application framework less demanding in terms of performance, power consumption, safety, etc. Developing the concept of a second life for components further contributes to sustainability by repurposing older hardware for new applications. Having hardware compatible with standard protocols (such as the one used for the web or cloud) could improve the chances for such a reuse.

2.1.6.3.2 Key focus areas

Ensuring Long-Term Functionality and Up-to-Date Operation

A fundamental aspect of supporting the increasing lifespan of devices is ensuring that they remain functional and up-to-date over extended periods. This involves not only maintaining their operational capabilities but also providing ongoing support for software updates and new features. By keeping devices updated with the latest advancements and security patches, manufacturers can extend their usability and relevance, thereby maximizing their value to users and reducing electronic waste. This may incur over-specification to be able to be future proof.

- Dynamic reconfiguration and Self-X techs

Intelligent reconfigurable concepts are an essential key technology for increasing the re-use and service life of hardware and software components. Such modular solutions on system level require the consideration of different quality or development stages of sensors, software, or AI solutions. If the resulting uncertainties (measurements, predictions, estimates by virtual sensors, etc.) are considered in networked control concepts, the interoperability of agents/objects of different generations can be designed in an optimal way.

- Dynamic reconfiguration: a critical feature of the AI circuits is to dynamically change their functions in real-time to match the computing needs of the software, AI algorithms and the data available, and create software-defined AI circuits and virtualise AI functions on different computing platforms. The use of reconfigurable computing technology for IoT devices with AI capabilities allows hardware architecture and functions to change with software providing scalability, flexibility, high performance, and low power consumption for the hardware. The reconfigurable computing architectures, integrated into AI-based circuits can support several AI algorithms (e.g. convolutional neural network (CNN), fully connected neural network, recursive neural network (RNN), etc.) and increase the accuracy, performance and energy efficiency of the algorithms that are integrated as part of software-defined functions.
- Realizing self-X (adaptation, reconfiguration, etc.): for embedded systems self-adaptation, self-reconfiguration has an enormous potential in many applications. Usually in self-reorganizing systems the major issue is how to self-reorganise while preserving the key parameters of a system (performance, power consumption, real-time constraints, etc.). For any system, there is an operating area which is defined in the multi-dimensional operating parameter space and coherent with the requirements. Of course, very often the real operating conditions are not covering the whole operating domain for which the system was initially designed. Thanks to AI, when some malfunctioning parts are identified, it could be possible to assess, relying on AI and the data accumulated during system operation, if it affects the behaviors of the system regarding its real operating conditions. If this is not the case, it could be considered that the system can continue to work, with maybe some limitations, but which are not vital regarding normal operation. It would then extend its lifetime “in place”. The second case is to better understand the degraded part of a system and then identify its new operating space. This can be used to decide how it could be integrated in another application making sure that the new operating space of the new part is compatible with the operating requirements of the new hosting system.
- Self-learning techniques are promising. Prediction on Natural Language Understanding (NLU) on the fly or keyboard typing, predictive maintenance on mechanical systems (e.g. motors) are more and more studied. Many domains can benefit of the AI in mobility, smart building, and communication infrastructure. LLMs shows interesting properties in this field, such as few-shot learning.

- Monitoring the devices and systems health

Dynamic reconfiguration and Self-X Techs also need data to be as efficient as possible. It means that monitoring of devices and systems health are necessary to acquire those data. Several techniques exist.

- Distributed monitoring: continuous monitoring and diagnosis also play a crucial role for the optimization of product lifetime. Where a large amount of data is collected during daily life operation (e.g. usage, environment, sensor data), big data analysis techniques can be used to predictively manipulate the operational strategy, e.g. to extend service life. Similarly, an increase in power efficiency can be achieved by adjusting the calibration in individual agents. For example, consider a fuel cell electric vehicle where the operation strategy decisively determines durability and service life. Distributed monitoring collects data from various interconnected agents in real-time (e.g. a truck platoon, an aircraft swarm, a smart electricity distribution network, a fleet of electric vehicles) and uses these data to draw conclusions about the state of the overall system (e.g. the state of health or state of function). This allows to detect shifting behavior or faulty conditions in the systems and to even isolate them by attributing causes to changes in individual agents in the network or even ageing of individual objects and components. Such detection should be accomplished by analyzing the continuous data stream that is available in the network of agents. A statistical or model-based comparison of the individual objects with each other provides additional insights. Thus, for example, early failures of individual systems could be predicted in advance. This monitoring should also cover the performance of the semiconductor devices themselves, especially to characterize and adjust to ageing and environmental effects and adjust operations accordingly.
- Distributed predictive optimization is possible, whenever information about future events in a complex system is available. Examples are load predictions in networked traffic control or demand forecasts in smart energy supply networks. In automation, a concept dual to control is monitoring and state observation, leading to safety-aware and reconfigurable automation systems. Naturally, all these concepts, as they concern complex distributed systems, must rely on the availability of vast data, which is commonly associated with the term big data. Note that in distributed systems the information content of big data is mostly processed, condensed, and evaluated locally thus relieving both communication and computational infrastructure.
- The other area of lifetime extension is how AI could identify very low signals in a noisy data environment. In the case of predictive maintenance, for instance, it is difficult for complex machinery to identify a potential failing part early in advance. The more complex the machinery, the less feasible it is to obtain a complete analytic view of the system, which would allow for simulation and thus identification of potential problems in advance. Thanks to AI and collecting large datasets it is possible to extract some very complex patterns which could allow for very early identification of parts with a potential problem. AI could not only identify these parts but also give some advice regarding when the exchange of a part is needed before failure, and then help in maintenance task planning.

- Artificial Intelligence challenges

Secure software upgradability is necessary in nearly all systems now and hardware should be able to support future updates. AI introduces additional constraints compared to previous systems. Multiplicity of AI approaches (Machine learning, DL, semantic, symbolic, etc.), multiplicity of neural network architectures based on a huge diversity of neuron types (CNN, RNN, LLMs, etc.), potential complete reconfiguration of neural networks for a same system (linked to a same use case) with a retraining phase based on an adapted set of data make upgradability much more complex. This this why HW/SW, related stacks, tools, data sets compatible with the edge AI system must be developed in synergy. HW/SW plasticity is necessary whatever the AI background principle of each system is, to make them as much as possible upgradable and interoperable and to extend the system lifetime. HW virtualization will help to achieve this, as well as standardisation. The key point is that lifespan extensions, like power management, are requirements which must be considered from day one of the design of the system. It is impossible to introduce them near the end without a strong rework.

One challenge for the AI edge model is upgradability of the firmware and new learning/training algorithms for edge devices. This includes the updates over-the-air and the device management of the updating of AI/ML algorithms based on the training and retraining of the networks (e.g. neural networks, etc.) that for IoT devices at the edge is very much distributed and is adapted to the various devices. The challenge of the AI edge inference model is to gather data for training to refine the inference model as there is no continuous feedback loop for providing this data. The related security questions regarding model confidentiality, data privacy etc. need to be addressed specifically for such fleets of devices.

The novelty with AI systems is to upgrade while preserving and guaranteeing the same level of safety and performance. For previous systems based on conventional algorithmic approaches, the behavior of the system could be evaluated offline in validating the upgrade with a predefined data set representative enough for the operating conditions, knowing that, more than the data themselves, the way they are processed is important. In the case of AI, things are completely different, as the way data are processed is not typically immediately understandable, but what is key is the data set itself and the results it produces. In these conditions it is important to have frameworks where people could reasonably validate their modification, whether it is hardware or software, in order to guarantee the adequate level of performance and safety, especially for systems which are human-life-critical. Another upgrade-related challenge is that of designing systems with a sufficient degree of architectural heterogeneity to cope with the performance demands of AI and machine learning algorithms, but at the same time to be flexible enough to adapt to the fast-moving constraints of AI algorithms. Whereas the design of a new embedded architecture or electronic device, even of moderate complexity, takes typically 1-3 years, AI models such as Deep Neural Networks are outdated in just months by new networks. Often, new AI models employ different algorithmic strategies from older ones, outdating fixed-function hardware accelerators and necessitating the design of hardware which functionality can be updated.

Designing with Modularity and Extensibility in Mind

To achieve long-term functionality, devices need to be designed with modularity and extensibility as core principles. Modularity allows to extend the functionality of a chip for example by externally adding another support chip. For this to be practically usable means that a specific chip-to-chip interface should be added and internally connected to ensure this modularity at board/system level. This design philosophy supports easier maintenance and enables incremental improvements, which can significantly extend the device's lifespan. Extensibility ensures that the device can adapt to new technologies and requirements over time at system level, adding extra features with new supporting chips. This seems at the cost of integration and performance, but new techniques could be developed that overcome most of the disadvantages. For example, Compression Attached Memory Module (CAMM)[53] still allows to replace the memories in a system like a PC, while the recent trend is to integrate memory directly on a (passive) interposer, thus preventing any memory upgrade.

- ### Advancing Modularity at the Chiplet Level

  While the concept of modularity is well-established at board and system level, achieving it at the level of chiplets is an emerging challenge that has not yet been fully realized. Chiplet-based architectures involve breaking down a monolithic System on Chip (SoC) into smaller, interconnected modules (chiplets) that can be individually upgraded or replaced. This approach promises greater flexibility, improved scalability, and easier integration of new technologies. However, the industry is still in the early stages of developing standardized, interoperable chiplet designs. Advancing this modularity at the chiplet level is crucial for enabling devices to support prolonged lifespans.

- ### Virtual Modularity

  Virtual Modularity refers to the dynamic integration and seamless interaction between multiple devices within an interconnected ecosystem, enabling them to function as a unified system. This concept allows different hardware components and devices to pool their resources and capabilities, creating a flexible, scalable environment where services and tasks can be effortlessly transferred or shared across the network. As a result, users experience a cohesive, uninterrupted operation, regardless of which device they are using, as the underlying system intelligently manages and allocates resources to optimize performance and deliver a consistent user experience across all connected devices[54].

  Virtual Modularity can increase the lifespan of devices by distributing workloads and optimizing resource usage across a network of connected devices, rather than relying on a single device to handle all tasks. Since different devices in the ecosystem can share and collaborate on tasks, the wear and tear on individual devices is reduced. For example, a smartphone may offload demanding processing tasks to a more powerful device, such as a tablet or computer, thereby reducing its own strain and energy consumption. This collaborative approach helps prevent any one device from being overworked, which can extend its operational life. Additionally, as new devices are added to the system, older devices can continue to contribute and function effectively within the network, further extending their useful life by leveraging the collective power and capabilities of the entire ecosystem.

---

[53] https://en.wikipedia.org/wiki/CAMM_(memory_module)

[54] This concept is developed for example in the Computing Continuum approach, and examples are the Continuity feature on Apple devices or the “super device” in HarmonyOS.

Older devices can have their lifetime increased thanks to the extra computing and storage resources delivered by other devices in the pool.

Reuse of components or systems in a downgraded or requalified use case

Re-use: One concept called the "2nd life" is actually the re-use of parts of systems. Such re-use could be adapted to edge AI as far as some basic rules are followed. First, it is possible to extract the edge AI HW/SW module which is performing a set of functions. For example, this module performs classification for images, movement detection, sound recognition, etc. Second, the edge AI module can be requalified and recertified downgrading its quality level. A module implemented in aeronautic systems could be reused in automotive or industrial applications. A module used in industrial could be reused in consumer applications. Third, an AI system may be re-trained[55] to fit the "2nd life" similar use case, going for example from smart manufacturing to smart home. Last, the business model will be affordable only if such "2nd life" use is on a significant volume scale. A specific edge AI embedded module integrated in tens of thousands of cars could be removed and transferred in a new consumer product being sold on the market.

2.1.6.4 Major Challenge 4: Ensuring European sustainability of embedded intelligence

2.1.6.4.1 State of the art

Edge computing and Artificial Intelligence are quickly becoming more and more tied together, converging towards the unified concept of Embedded Artificial Intelligence. To ensure the growth of this emerging technologies it is crucial that the European ecosystem of companies and academia can strategically cover all the steps of the associated value chain. This means covering the entire stack starting from the Embedded Artificial Intelligence hardware, including all the software layers required to exploit the hardware functionalities and to develop the final applications, the engineering of the tools for AI development and the data sets, with a trustable and certifiable environment.

Economic and environmental sustainability represents to key factors when defining actions and investments promoting the coverage of the value chain and of the technology stack. Technology is strongly affected by sustainability that, very often, tips the scale between the ones that are promising, but not practically usable, and the ones making the difference. E.g. cloud computing, based on data centers, plays a fundamental role for the digitalization process. However, data centers consume a lot of resources (energy[56], water, etc.), they are responsible for significant carbon emissions during their entire lifecycle and generate a lot of electronic and chemical waste.

[55] In the domain of AI using LLMs, the extension of lifetime of foundational models could be performed by several approaches, e.g by evolution, i.e. fine-tuning of the model. It might also be possible to "retune" the use of the LLMs in the latent space, i.e. without changing the parameters of the networks, but only by changing the context of the "prompts". This needs further analysis for being effectively applicable for practical reuse.

[56] Andrae, Anders. (2017). Total Consumer Power Consumption Forecast

Today, the percentage of worldwide electricity consumed by datacenters and data transmission networks is 2 to 3%. "*Rapid improvements in energy efficiency have, however, helped moderate growth in energy demand. Data centres and data transmission networks are responsible for 1% of energy-related GHG emissions*" [57]. A report from January 2024 of IEA[58] indicates that "Global electricity demand from data centers could double towards 2026" (see Figure 2.1.15). Interestingly, the electricity demand to process crypto-currencies seems still to surpass the consumption related to AI in 2026 (see Figure 2.1.16).

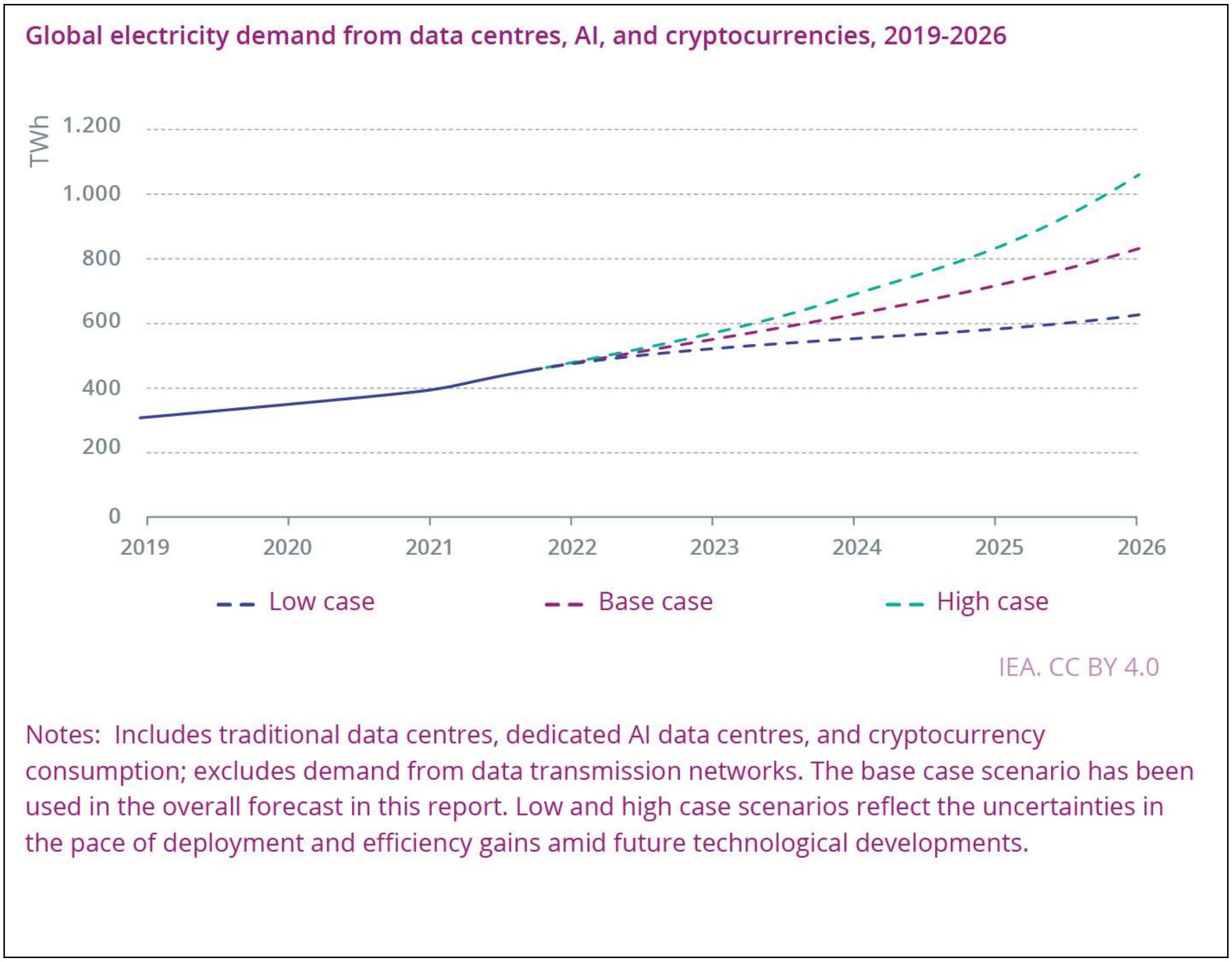


Figure 2.1.15: Forecast of growth of global electricity demand from data centers, from [72]

[57] https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks#tracking, captured in August 2024

[58] https://iea.blob.core.windows.net/assets/6b2fd954-2017-408e-bf08-952fdd62118a/Electricity2024-Analysisandforecastto2026.pdf

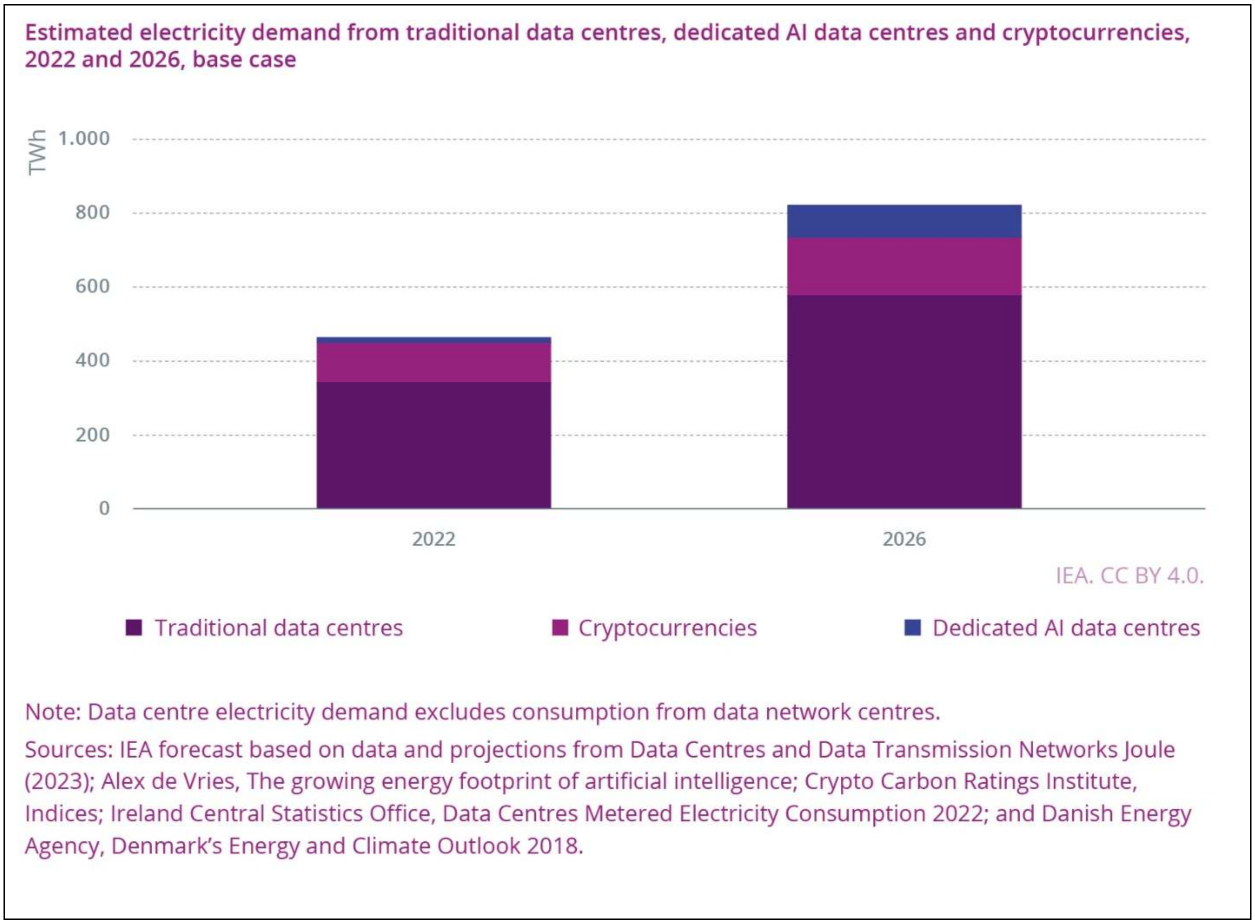


Figure 2.1.16: growth of cryptocurrency and AI centric data centers between 2022 and 2026, from [72]

Data centers are progressively becoming more efficient, but shifting the computing to the edge, for example, allows to temporally reduce data traffic, data centers storage and processing. However, only a new computing paradigm could significantly reduce their environmental footprint and ensure sustainability. Edge Computing could contribute to reach this goal by the introduction of ultra-low power and efficient computing solutions, and Embedded Intelligence (and collaborative AI agents approach) can significantly contribute to optimize their operations and, directly and indirectly, their overall energy consumption profile.

Shifting to green energy is certainly a complementary approach to ensure sustainability, but the conjunction of AI and edge computing, the Embedded AI, has the potential to provide intrinsically sustainable solutions with a wider and more consolidated impact. Indeed, a more effective and longer-term approach to sustainable digitalization implies reconsidering the current models adopted for data storage, filtering, analysis, processing (including AI training itself), and communication. By embracing Edge Computing, for example, it is possible to significantly reduce the amount of useless and wasteful data flowing to and from the cloud and data centers, because it means adopting architectural and structural solutions that are more efficient and that permanently reduce the overall power consumption. These solutions also bring other important benefits such as real-time data analysis, reduced memory and storage capacity and better data protection. The Edge Computing paradigm also makes AI more sustainable: it is evident that cloud-based machine learning inference is characterized by a huge network load, with a serious impact on power consumption and huge costs for organizations. Transferring machine learning inference and data pruning to the edge, for example, could exponentially decrease the digitization costs and enable sustainable businesses. To avoid this type of drawbacks, new AI algorithms specifically tailored for embedded systems, in conjunction with new HW components (e.g. based on neuromorphic architectures), should be designed and developed. Considering the application areas, in some cases, this could lead to more specialized and extremely efficient solutions.

Sustainability of Edge Computing and AI is affected by many technological factors, which require investments to consolidate and strengthen EU global positioning, generating a potential positive fallback on the sustainability of future digitalization solutions and related applications. GAMAM already master these technologies and are already controlling the complete value chain and technology stack associated with them: to achieve strategic autonomy in this field, Europe must cover the entire value chain, consolidating the existing segments, identifying/filling the technology gaps, and leveraging on a strong cooperation between the European stakeholders, with a particular attention to SMEs (which generate a large part of European GDP). From this perspective, European coordination to develop AI, edge computing and Embedded AI technologies is fundamental to create a sustainable value chain, based on solid alliances, and capable to support the European key vertical applications.

Ensuring sustainability for Embedded AI in Europe involves several key initiatives: open-source hardware can drive innovation and reduce costs, while energy efficiency improvements are critical for sustainable development. Engineering support represents the third key factor and is necessary to advance sustainable Embedded AI from several perspectives: optimize the energy profile of Embedded AI solutions (technical and environmental levels), tackle the lack of expertise and of human resources (operational/engineering level), and keep the costs of these solutions sustainable (economic). For example, by mastering the adoption of new Embedded AI techniques, such as transformers and large language models (LLMs) with a focus on federation of smaller specialized models developed in Europe[59], and integrating them with efficient accelerators and algorithms, Europe can maintain a competitive edge. This is crucial for many vertical domains that are key for Europe: for example, in the domain of autonomous systems, Europe has the potential to lead in the development of autonomous vehicles and robots, leveraging advanced AI and embedded intelligence. The necessity to ensure an efficient engineering process, with consequent short time-to-market, is motivated by the speed of global market evolution: for example, the emergence of cheap Chinese RISC-V microcontrollers emphasizes the necessity for fast development, robust, cost-effective computing solutions made in Europe. Another example from a vertical application perspective is the partitioning of complex systems on a chip (SoC) into chiplets and interposers, which requires the development of an ecosystem of interoperable chiplets, comprehensive architecture and engineering solutions to support this approach: the automotive market can significantly benefit from this, maintaining Europe at the global market forefront.

Sustainability can also be measured in terms of democratizing technology use, to make it more accessible to a broader community of users and facilitate market entry to startups and SMEs: natural language interfaces for example could contribute to facilitate the access to digital technologies. This requires that the associated AI models are trained with European data covering European languages, respecting European ethics. Additionally, democratizing technology also means supporting the diversity of European computing systems/chips, simplifying their programmability and consolidating this diversity through

---

59 For example, see https://mistral.ai/

interoperability. This is also crucial to support the research into new computing paradigms, like neuromorphic and analog computing, and to ensure they can reach high TRL levels (i.e. 7, 8 and 9).

One of the major challenges that need to be accounted for in the next few years is related to the design of progressively more complex electronic systems to support advanced functionalities such as AI and cognitive functionality, especially on the edge. This is particularly challenging in the European landscape, which is dominated by small and medium enterprises (SMEs) with only some large actors that can invest and support larger-scale projects. To consolidate and strengthen European competitiveness and ensure sustainability in advanced Embedded AI solutions it is therefore crucial to create an ecosystem covering the entire value chain, in which SMEs can cooperate and increase their level of innovation and productivity. The definition of open industrial standards and a market of Intellectual Properties (IPs) is crucial to accelerate the design and development of these solutions, guaranteeing EU competitiveness and increasing the market dimensions. Open-source software, hardware and engineering tools can play an extremely important role in this regard, because they allow to reduce engineering costs for licensing and verification significantly, lowering the entry barrier to design innovative products.

2.1.6.4.2 Key focus areas

Leveraging Open-source Hardware and Software for Innovation and Cost Reduction

Open-source Hardware and Software are drivers of innovation and cost reduction in the fields of Edge Computing and Embedded AI, allowing democratization of design and creating a collaborative environment where innovation can emerge and making the development of cutting-edge solutions that are both affordable and highly effective possible. Open-source also facilitates transparency and security, which are essential factors for building trust and compliance with European standards.

Developing and Federating Smaller Specialized AI Models

The federation of smaller specialized AI models developed in Europe represents a strategic approach to harnessing localized expertise and enhancing the efficiency of AI applications at the edge. These models, when federated (in "*Agentic AI*"), can work together to tackle complex tasks as, or even more effectively than monolithic models. Promoting the development of these models in Europe contributes to the alignment of technology with European technologies and applications necessities, leveraging regional strengths in AI research and development. Specialized agents can be derived from few foundation models (made by few European companies or from Open Source) by fine tuning that can be achieved by SMEs or start-ups, as this specialization requires less computing power and data than the creation of a foundation model.

Training Models with European Data and Ethical Compliance

For AI models to be truly sustainable and beneficial to Europe, they must be trained on data that covers European languages and reflects the diverse cultural and social contexts of the continent. This ensures that AI applications are more accurate, inclusive, and useful across different European regions. Additionally, adherence to European ethical standards in AI development is crucial. This can be achieved:

- Finding solutions which ensure privacy, fairness, and transparency in AI systems, thereby aligning technological advancements with the core values and regulations of Europe.

- Allowing those AI to work at the edge, hence promoting the development of specialized AI chips for edge AI, this will facilitate access to digital technologies to a broad audience.
- Supporting natural language interfaces which enable users to interact with technology in intuitive ways, breaking down barriers to adoption and use, in a continent characterized by cultural, societal and language diversity.
- Increasing the re-use and sharing of knowledge and models generated by embedded intelligence, improving the energy- and cost-efficiency of AI training, and adopting a new benchmarking AI approach oriented to sustainability.

Deploying Efficient and Sustainable Embedded AI-oriented ECS

Efficient ECS like new memories, accelerators and algorithms specifically conceived for Embedded AI are fundamental to the performance and sustainability of the domain. Developing and deploying these technologies within Europe ensures that AI applications can operate with high efficiency, reducing energy consumption and enhancing overall system performance. Most current designs involve perception-based AI accelerators, but Europe should be at the forefront to use new disruptive technologies such as generative AI and ensuring its sovereignty by developing the HW and SW that will enable it at the edge. This will require for example:

- New materials and substrates oriented to low power consumption.
- New generations of embedded ultra-low power electronic components.
- 3D-based device scaling for low power consumption and high level of integration.
- Chiplet-based solutions promoting modularity and reuse (sustainability by design).
- Strategies for self-powering nodes/systems on the edge and efficient cooling solutions.
- Policies and operational algorithms for power consumption at edge computing level.
- Efficient and secure code mobility.
- Advanced Neuromorphic components.
- Generative AI for the edge.
- Inclusion of existing embedded systems on the edge (huge market opportunity).

Accelerating Development of Robust, Cost-Effective Solutions

There is a pressing need for the rapid development of robust, cost-effective computing solutions made in Europe. These solutions must be designed to meet the specific demands of European markets and regulatory environments. By focusing on speed of realization, reliability, and cost-effectiveness, European manufacturers can remain competitive and responsive to global and local market needs. This involves developing the full spectrum of technologies, from hardware innovations to sophisticated software solutions, to new tools allowing to increase productivity, ensuring that European products stand out in the global marketplace. For example, this must include:

- Engineering process automation for full lifecycle support of Embedded AI solutions.
- Edge Computing and Embedded AI security and sustainability by design.
- Engineering support for Embedded AI verification and certification, addressing end-to-end edge solutions.
- Support for education and professional training to enable, facilitate, improve and speed-up the design, development and deploy of Embedded AI.
- Engineering process automation based on generative AI.

2.1.7 Timeline

As the field concerning Artificial Intelligence is evolving very rapidly, all predictions are subjective, especially for 2035 and beyond (the foundation paper about Transformers, which lead to the current LLMs, was only published in 2017[60], so in 2015 nobody could have predicted what we are seeing today in this field). It is why the long-term column (2035 and beyond) is rather empty.

Legend:

- (EC): edge computing
- (eAI:) Embedded Artificial Intelligence

[60] https://arxiv.org/abs/1706.03762

MAJOR CHALLENGE	TOPIC	SHORT TERM	MEDIUM TERM	LONG TERM
		2025-2029	2030-2034	2035 AND BEYOND
Major Challenge 1: increasing the energy efficiency of computing systems	**Processing data locally and reducing data movements: towards the computing continuum**	Move towards the continuum of computing: orchestrators can select computing where it is the most efficient (statically). Using the same software development infrastructure from deep edge to edge and possibly HPC applications. Use of similar building blocks from deep edge to edge devices. Developing open architectures (for fast development) with maximum reuse of tools and frameworks Development of edge (ex: fog) type of computing (peer to peer), development of distributed systems e.g. for federated learning or fine tuning of LLMs reducing need to exchange data. Unified memory: avoid copying data from the CPU memory to the accelerator(s) memory(ies).	Advance storage management: towards distributed storage supporting the continuum of computing approach also for storage. Move towards the continuum of computing: orchestrators can select or move computing where it is the most efficient (dynamically) avoiding data movements.	
	Co-design of algorithms, hardware, software	AI can be used to help this co-design. This is mainly a topic for the "methodology and tools" chapter. Tools allowing fast realization of hardware accelerators when a new paradigm emerges (e.g. Transformers/generative AI). Tools allowing semi-automatic design exploration of the space of configurations, including variants of algorithms, computing paradigms, hardware performances, etc. Complete 2.5D (interposers and chiplets) ecosystem, with tools increasing productivity and reuse of chiplets in different designs Automatic adaptation of complex neural networks or emerging AI algorithms to embedded systems with a minimum loss of performance	Systems can generate hardware (orchestrators, accelerators, storage and communication) and the associated low-level software from high-level specifications. Auto-configuration of a distributed set of resources to satisfy the application requirements (functional and non-functional) Supporting tools integrating multiple computing paradigms in the same package.	

	Efficient management of storage resources	Compressing weights (from FP32 to INT4 or less), pruning neural networks. Federations of multiples devices into "meta-devices" that share storage resources. Dynamic use of adapters/LoRA to switch the specialization of a foundation model (changing a small proportion of its weights[61]). Easily upgradable LLM accelerators, fine tuning "on premises". Innovation in memory technology: Using directly the parameters of neural networks from non-volatile memory without transferring them to RAM, computing near/in memory Create gateways between various solutions, beyond ONNX (for eAI)	New memory technology with least cost, fast access and certain level of non-volatility allowing to redefine the memory hierarchy completely.	
	Energy proportionality	Scalable accelerators where parts can be switched off. Refining dynamic power management techniques to ensure that energy consumption closely matches the workload, minimizing power wastage during low activity periods. Advancements in adaptive voltage scaling, power gating, and more efficient power distribution networks.	Use AI to dynamically manage the repartition of active parts in a chip.	
	Ultra-low standby current	Improving energy efficiency at low energy level, reducing leakage currents in devices, optimizing power management at the software level. Develop always-on units that can react to information in order to wake-up the rest of the system.	Pushes further towards near-zero standby power. Overcoming the limitations of current materials. Developing new semiconductor technologies that can operate at ultra-low power levels. Integrating advanced energy harvesting techniques to sustain device functionality even in standby mode.	

[61] It seems that Apple is doing something similar in its « Apple Intelligence »

	Development of innovative hardware architectures	Development of computing paradigms (e.g. using physics to perform computing, e.g. neuromorphic). Use of other technologies than silicon (e.g. photonics) Use of 2.5D, interposers and chiplets, with efficient interconnection network, e.g. using photonics) Creating an ecosystem around interposers and chiplets, with interoperability standards New In-memory computing accelerators Supporting software for all these hardware innovations	Dynamic instantiation of multi-paradigm computing resources according to the specifications of the task to be performed. New (de facto) standard allowing automatic interfacing, discovery, and configuration of resources (computing, storage) Global reconfiguration of the resources to satisfy the functional and non-functional requirements (latency, energy, etc.) Development of hybrid architectures, with smooth integration of various processing paradigms (classical, neuromorphic, deep learning), including new OSs supporting distributed computing of multiple computing paradigms Advanced In-memory computing accelerators	Integration in the same package of multiple computing paradigms (classical, Deep Learning, neuromorphic, photonic, quantum, etc.) Exploring potential use of quantum computing in Artificial Intelligence?

Major Challenge 2: managing the increasing complexity of systems	**Balanced mechanisms between performance and interoperability**	Exposing the non-functional characteristic of devices/blocks and off-line optimization when combining the devices/blocks Explore AI techniques for Self-x	On-line (dynamic) reconfiguration of the system to fulfil the requirements that can dynamically change (Self-x) Use of AI techniques for Self-x.	Drive partitioning through standards
	Developing distributed edge computing systems	Introduce standard APIs comprising Web based technologies; support of a layered AI architecture See items above in *Increasing the energy efficiency of computing systems*	Development by AI, e.g. of communication and functional collaboration, orchestration. Agentic (AI) approach See items above in *Increasing the energy efficiency of computing systems*	See items above in *Increasing the energy efficiency of computing systems*

	Scalable and Modular AI	Decomposition of SoCs, e.g., into chiplets See items above in *Increasing the energy efficiency of computing systems*	Ecosystem of chiplets for embedded control and embedded AI established Data and learning driven circuits design See also items above in *Increasing the energy efficiency of computing systems*	See items above in *Increasing the energy efficiency of computing systems*
	Easy adaptation of models	Development of efficient and automated transfer learning: only partial relearning required to adapt to a new application (Ex: Federative learning) Create a European training reference database for same class of applications/use cases network learning	Optimization of the Neural Network topology from a generically learned network to an application specific one.	Generic model-based digital AI development system
	Easy adaptation of modules	Easy migration of application on different computing platforms (different CPU – x86, ARM, RISC-V, different accelerators)	Use of HW virtualization Automatic transcoding of application for a particular hardware instance (à la Rosetta 2)	Generic model based digital development system
	Realizing self-X **Self-optimize, reconfiguration and self-management**	Add self-assessment features to edge devices Explore what AI techniques (such as LLMs) can do	Automatic reconfiguration of operational resources following the self-assessment to fulfil the goal in the most efficient way Deploy AI based approaches for self-optimization	Modelling simulation tools for scalable digital twins
	Using AI techniques to help in complexity management	Using AI techniques for the assessment of solutions and decrease the design space exploration	Automatic generation of architecture according to a certain set of requirements (in a specific domain)	Modelling simulation tools for scalable digital twins

	Using AI techniques modeling of interactions among system components	Explore usage of AI-based methods, tools and environments for the modeling and simulating of the system designs and their components	Leveraging networks of agents (LLMs and other AI-based programs) for automatized modeling and simulations	

Major Challenge 3: supporting the increasing lifespan of devices and systems	**HW supporting software upgradability** **(eAI)**	Create European training reference databases for same class of applications/use cases network learning Develop European training benchmarks (Methods and methodologies) Build framework tools for HW/SW for fast validation and qualification Establish interfaces standards compatible with most of AI approaches	HW virtualization based on AI algorithms Generic AI functions virtualization European training standards (Compliance/Certification) Certifiable AI (and paths towards explainability and interpretability) Interoperability and extensibility within computing continuum	Explainable AI Universal control
	Realizing self-X Also partially in ***Managing the increasing complexity of systems*** ***(eAI)***	Unsupervised learning technics Development of efficient and automated transfer learning: only partial relearning required to adapt to a new application (Ex: Federative learning)	HW virtualization based on AI algorithms Generic AI functions virtualization Certifiable AI (and paths towards explainability and interpretability) Use of Mixed of AI Agents and/or Transformers for generative AI on self-X Techs	Explainable AI Universal control

Improving interoperability (with the same class of application) and between classes, modularity, and complementarity between generations of devices. **(EC)** Also, partially in ***Increasing the energy efficiency of computing systems***	Developing open architectures (to quickly develop) with maximum reuse of tools and frameworks Interfaces standards (more than solutions) (could help explainability move from black to grey boxes)	Generic functions modules by class of applications/use cases + virtualization Use of Mixed of AI Agents and/or Transformers for generative AI Chiplets Virtual Modularity	
Improving interoperability of AI functions (with the same class of application) and between classes, modularity, and complementarity between generations of devices. **(eAI)** Also, partially in ***Increasing the energy efficiency of computing systems***	Developing open AI architectures (to fast develop) with maximum reuse of tools and frameworks Interfaces standards (more than solutions) (could help explainability of AI with a move from black to grey boxes) Clarified requirements for embedded AI in industry Applications and Apps working simultaneously through different equipment	Generic AI functions modules by class of applications/use cases + virtualization	Universal control

	Developing the concept of 2nd life for components **(EC)** (Link with sustainability)	Inclusion of existing embedded systems on the edge (huge market opportunity)	Generic set of functions for multi-applications/use cases Library of generic set of functions (Standardization) Basic data collection for predictive maintenance Global data collections for predictive maintenance by applications/use cases	Standardize flow for HW/SW qualification of generic set of functions (including re-training) which are used in a downgraded application/use case Full chain of reuse / Ecodesign conception

Major Challenge 4: ensuring European sustainability of embedded intelligence	**Leveraging Open-source Hardware and Software for Innovation and Cost Reduction**	Open-source eAI software Open-source training datasets for eAI Open-source foundation models for eAI	Open-source eAI hardware	
	Developing and Federating Smaller Specialized AI Models	AI specialized models tailored for eAI eAI models for specific applications eAI models modularity and reuse Partial eAI models federation	Interdisciplinary eAI models eAI models covering cross-domain applications Extensive eAI models federation	
	Training Models with European Data and Ethical Compliance	European-native training data sets Methods and HW/SW solutions respecting privacy, fairness, and transparency Natural language support in embedded systems Energy- and cost-efficiency of eAI training Specialized eAI chip to facilitate market entry	European-native training data sets Methods and HW/SW solutions respecting privacy, fairness, and transparency Natural language support in embedded systems Energy- and cost-efficiency of eAI training	

Deploying Efficient and Sustainable Embedded AI-oriented ECS	Materials and electronic components oriented to low and ultralow power solutions Strategies for self-powering nodes/systems on the edge Efficient cooling solutions Chiplet-based solutions for modularity and reuse (sustainability by design) Inclusion of legacy embedded systems Neuromorphic components Integrated power consumption management	3D-based device scaling for low energy consumption Efficient and secure code mobility. Advanced Neuromorphic components.	
Accelerating Development of Robust, Cost-Effective Solutions	Engineering process automation Continuous engineering across the product life cycle eAI security and sustainability by design eAI HW and SW modularity Adoption of generative AI in the eAI engineering process	Holistic development environment for eAI Engineering support for verification and certification eAI security and sustainability by design Reuse of knowledge and models generated by embedded intelligence	

2.2

Cross-Sectional Technologies

CONNECTIVITY

2.2. Connectivity

2.2.1 SCOPE

Connectivity and interoperability technologies in ECS enabling business and social benefits are tied to layers 1, 5, and 6 of the OSI model. The focus on these layers is motivated by the Major Challenges that characterize them (see Figure 2.2.1).

2.2.1.1 Scope for OSI layer 1

The scope covers the following types of physical layer connectivity.

- Cellular:
  - Beyond 5G
  - 6G
  - Direct to cell (Supplemental Coverage from Space scenario)

- Satellite communication
  - Low Earth Orbit and geosynchronous equatorial orbits
  - Non Terestrial Network

- Low power wide area:
  - Cellular: narrow band IoT, LTE, 6G, etc.
  - Non-cellular: SigFox, LoRa, M-Bus, etc.

- Low power short range:
  - Wireless: existing (Bluetooth, WiFi, etc.) or innovative technologies (mmW, etc.).
  - Wired: covering both high-speed optical glass and plastic fibers, mmW plastic fibers, and copper interconnect (USB, DOCIS, PCIe, UCIe, AIB, etc.).

- High speed:
  - Wireless: point to point mmW .
  - Wired: high-speed optical (800 Gb+, etc.) and copper interconnect (Ethernet, etc.).

The main challenge will be to ensure European leadership in terms of connectivity technologies (for example, standards) as well as associated software and hardware technologies supporting the development of connectivity solutions (e.g. chipset, module, protocols, etc.).

2.2.1.2 Scope for OSI layer 5 and 6

The scope addressed in this context is the interoperability from application to application relying on technologies at OSI layers 5 and 6. This interoperability covers the following underlying aspects:

- Protocols at all technology levels: Internet, operational and legacy.
- Security: such as protocol security, payload encryption, authentication, authorization, certificates, tokens and key distribution.
- Data semantics: supporting application to application understanding of transferred data/information.

LAYER		DATA UNIT	FUNTION
HOST LAYERS	7. Application		Network process to application.
	6. Presentation	Data	Data representation, encryption and decryption, convert machine-dependent data to machine-independent data.
	5. Session		Interhost comunication, managing sessions between applications.
	4. Transport	Segments	Reliable delivery of segments between points on a network.
MEDIA LAYERS	3. Network	Packet/Datagram	Addressing, routing and (not necessarily reliable) delivery of datagrams between points on a network.
	2. Data link	Bit/Frame	A reliable direct point-to-point data connection.
	1. Physical	Bit	A (not necessariliy reliable) direct point-to-point data connection.

Figure 2.2.1 - Major Challenges: OSI Model

PHYSICAL LAYER CONNECTIVITY	ECS KEY APPLICATIONS					
	MOBILITY 3.1	ENERGY 3.2	DIGITAL INDUSTRY 3.3	HEALTH AND WELLBEING 3.4	AGRIFOOD AND NATURAL RESOURCES 3.5	DIGITAL SOCIETY 3.6
Cellular	X	X	X	X	X	X
Low power wide area		X	X	X	X	X
Low power short range	X		X	X		X
High speed	X	X	X			X

Figure 2.2.2 - Major Challenge: ensuring European leadership in terms of connectivity technologies

2.2.2 Applications breakthroughs

Improvements in connectivity technology will have an impact on all ECS application areas. For health and well-being, connectivity interoperability issues are addressed by enabling faster translation of ideas into economically viable solutions, which can be further scaled up in daily health practice. Examples of health and well-being application breakthroughs supported here are:

- A shift in focus from acute, hospital-based care to early prevention.
- Strengthening where and how healthcare is delivered, supporting home-based care.
- Stronger participation of citizens in their own care processes, enhancing patient engagement.
- Supporting the clinical workforce and healthcare consumers to embrace technology-enabled care.
- Data communication technology for interoperability of wireless data infrastructure.

Heterogeneous and chiplet integration connectivity enabling European research, prototyping and large scale production addressing market needs in key application domains like e.g. automotive (SDV/Risc-V), HPC, industrial.

Improved, secure and interoperable connectivity will further support healthcare and well-being application breakthroughs regarding, for example:

- Healthcare deployment, enabling digital health platforms.
- Healthcare system paradigm transition from treatment to health prevention, enabling the shift to value-based healthcare.
- Building a more integrated care delivery system, supporting the development of the home as the central location for the patient.
- Enhancing access to personalized and participative treatments for chronic and lifestyle- related diseases.
- Enabling more healthy life years for an ageing population.

In the mobility application area, the provision of improved, robust, secure and interoperable connectivity will support breakthroughs regarding:

- Achieving the Green Deal for mobility, with the 2Zero goals of –37.5% CO_2 by 2030.
- Increasing road safety through the CCAM program.
- Strengthening the competitiveness of the European industrial mobility digitization value chain.

In the energy application domain, the provision of improved, robust, secure and interoperable connectivity will support breakthroughs regarding:

- Significant reduction of connectivity energy demand.
- Enabling necessary connectivity to the integration of the future heterogeneous energy grid landscape.
- "Plug and play integration" of ECS into self-organized grids and multi-modal systems.
- Solving safety and security issues of self-organized grids and multi-modal systems.

In the industry application domain, the provision of improved, robust, secure and interoperable connectivity will support closing gaps such as:

- Preparing for the 5G era in communications technology, especially its manufacturing and engineering dimension.
- Long-range communication technologies, optimized for machine-to-machine (M2M) communication, a large number of devices and low bit rates, are key elements in smart farming.
- Solving IoT cybersecurity and safety problems, attestation, security-by-design, as only safe, secure and trusted platforms will survive in the industry.
- Interoperability-by-design at the component, semantic and application levels.

- IoT configuration and orchestration management allowing for the (semi)autonomous deployment and operation of large numbers of devices.

In the digital society application domain, the provision of improved, robust, secure and interoperable connectivity will support the overall strategy regarding:

- Enabling workforce efficiency regardless of location.
- Stimulating social resilience in the various member states, providing citizens with a better work/life balance and giving them freedom to also have leisure time at different locations.
- Ubiquitous connectivity, giving people a broader employability and better protection against social or economic exclusion.
- Enabling European governments, companies and citizens to closer cooperation, and to develop reliable societal emergency infrastructures.

In the agrifood application domain, the provision of improved, robust, secure and interoperable connectivity will support innovations addressing the EU Green Deal regarding:

- Reducing the environmental impact related to transport, storage, packaging and food waste.
- Reducing water pollution and greenhouse gas emission, including methane and nitrous oxide.
- Reducing the European cumulated carbon and cropland footprint by 20% over the next 20 years, while improving climatic resilience of European agriculture and stopping biodiversity erosion.

2.2.3 MAJOR CHALLENGES

Five Major Challenges have been identified in the connectivity domain:

- **Major Challenge 1**: strengthening the EU connectivity technology portfolio to maintain leadership, secure sovereignty and offer an independent supply chain.
- **Major Challenge 2:** investigate innovative connectivity technology (new spectrum or medium) and new approaches to improving existing connectivity technology to maintain the EU's long-term leadership.
- **Major Challenge 3**: autonomous interoperability translation for communication protocol, data encoding, compression, security and information semantics.
- **Major Challenge 4**: architectures and reference implementations of interoperable, secure, scalable, smart and evolvable IoT and SoS connectivity from edge to cloud
- **Major Challenge 5**: network virtualization enabling run-time and evolvable integration, deployment and management of edge to cloud network architectures.

2.2.3.1 Major Challenge 1: Strengthening the EU connectivity technology portfolio to maintain leadership, secure sovereignty and offer an independent supply chain

2.2.3.1.1 State of the art

Today's connectivity solutions require an incredibly complex electronic system comprising various functions integrated into a wide range of technologies.

Note that advanced digital functions such as the application processor and the baseband modem are mastered by a limited number of US and Asian players (Mediatek, Qualcomm and Samsung), and achieved in advanced complementary metal–oxide–semiconductor (CMOS) technology available at only two Asian businesses (Taiwan Semiconductor Manufacturing Company, TSMC, and Samsung). On this last point, it is worth noting that through the America CHIPS Act, Intel, SAMSUNG and TSMC will build new advanced logic (3 nm and below) 300 mm Fabs in Arizona to strengthen US sovereignty and supply chain.

The European Chips Act supports ST and GF new 300 mm facility in France build to allow for 18 nm and beyond FDSOI technology manufacturing (while <10 nm R&D activities are pursued both at CEA Leti and IMEC), all this in order to limit the reliance on Asian foundries' manufacturing capabilities. Consequently for the foreseeable future, Europe most advanced semiconductor manufacturing capability will be limited to 18/16 nm CMOS node which a consequence of the products currently designed by European semiconductor players and required by associated key verticals (automotive, industrial, ...). This point is also illustrated by the decision of TSMC to start building during summer 2024 a new 16/12 nm 300 mm Fab in Dresden (11B $ investment) through the creation of a joint venture working with Infineon, NXP, and Bosch (each holds a 10% stake)

In fact, European players (Infineon, NXP, ST, etc.) are strong on the analogue and RF front end module markets, mainly due to the availability of differentiated technologies developed and manufactured in Europe. Differentiated technologies are a key strength of the European ECS industry, especially when considering the connectivity market. Consequently, to maintain Europe's leadership and competitiveness it is vital to ensure that European differentiated semiconductor technologies remain as advanced as possible. This is key to ensure that Europe secures the market share in the connectivity market, and also strengthens its technology leadership by playing a major role in the development and standardization of future connectivity technologies. This point is crucial to secure Europe's sovereignty on the connectivity topic.

Moreover, over the last year the rising economic tension between the US and China has underlined the value of Europe's ECS supply chain. Once again, this is especially true for differentiated technologies. For example, advanced BiCMOS technologies are currently mastered by a limited number of US (GlobalFoundries and TowerJazz) and European (Infineon, ST and NXP) players. With Chinese companies being forced to move away from US providers, this creates a significant opportunity for Europe as the only viable alternative. Consequently, strengthening Europe's connectivity technology portfolio and associated manufacturing capacity to offer an independent and reliable supply chain is now a key challenge for all European ECS actors.

In addition to being able to provide the differentiated semiconductor technologies supporting the development of innovative connectivity solutions, it is important to note that some European players are proposing connectivity chipset solutions (for example, Sequans Communications and Nordic on the narrowband IoT topic) or full connectivity solutions. Supporting the growth of these existing actors and helping emerging industry leaders is also a key challenge for Europe to capture a bigger proportion of the value chain, as well as to ensure its sovereignty on the connectivity topic in the long run.

2.2.3.1.2 VISION AND EXPECTED OUTCOME

To address identified connectivity technology challenges, we propose the vision described below, which can be summarized by the following three key points (with associated expected outcomes).

Strengthening Europe's differentiated technologies portfolio

As discussed above, Europe's differentiated semiconductor technologies are key assets that should be both preserved and improved upon to secure European leadership in connectivity. Consequently, dedicated research should be encouraged, such as the technologies below (which are also promoted in the Chapter 1.1 on Process Technology, Equipment, Materials and Manufacturing):

- ***Advanced BiCMOS***: targeting RF and sub-THz (i.e., 100-300GHz) and THz front-end modules.
- ***RF SOI***: targeting <7 GHz and mmW front-end modules.
- ***GaN and Gan on Si***: targeting the high-power infrastructure and high efficiency/wide bandwidth 6G handset markets (5.925 GHz – 7.125 GHz and 10 GHz – 13.25 GHz bands).
- ***FD SOI***: targeting power-efficient connectivity solutions (for example NB IOT and NTN).
- ***GaAs***: targeting mmW space and defense applications (W & K bands).
- ***InP & InP on Si***: targeting high-speed optical link (>800 Gb/s), mmW applications (6G sub-THz communication >100 GHz) and ultra-low noise Front End Module in K band.
- ***Silicon Photonics***: targeting next generation silicon photonics technology (SiN waveguide, thin film LNO, ...) to address both pluggable optics for data centre and AI driven co-package optics solution
- ***RF filters***: supporting the development on innovative European based technology (for example POI based TF SAW technologies).

- ***Advanced packaging***: enabling prototyping and medium volume production capability in Europe to enable heterogeneous integration of differentiated technology manufacture in Europe to move higher in the value chain and capture more value.

The main challenge will be in to improve achievable performances. To illustrate this, we have extracted the medium-term (2025) and long-term (2030) solid state technology roadmap proposed by H2020 CSA project NEREID to serve as a connectivity roadmap (see Figure 2.2.3). We can see that whatever the type of application (device-to-device, D2D, indoor, outdoor), the requirements in analogue RF will mainly consist of achieving Fmax and FT ~500 GHz in 2025 and 1 THz in 2030, while NFmin will be well below 1 dB in the medium term, to reach 0.5dB in the long term. The only parameter that differentiates the types of applications is the output power, which outdoors should reach between 36 and 40 dBm per PA by the end of the decade. The biggest challenge for silicon or hybrid-on-silicon substrate technologies is expected to be the frequency challenge. Technologies such as GaN/Si and RF SOI will deliver power but for applications operating at less than 100 GHz.

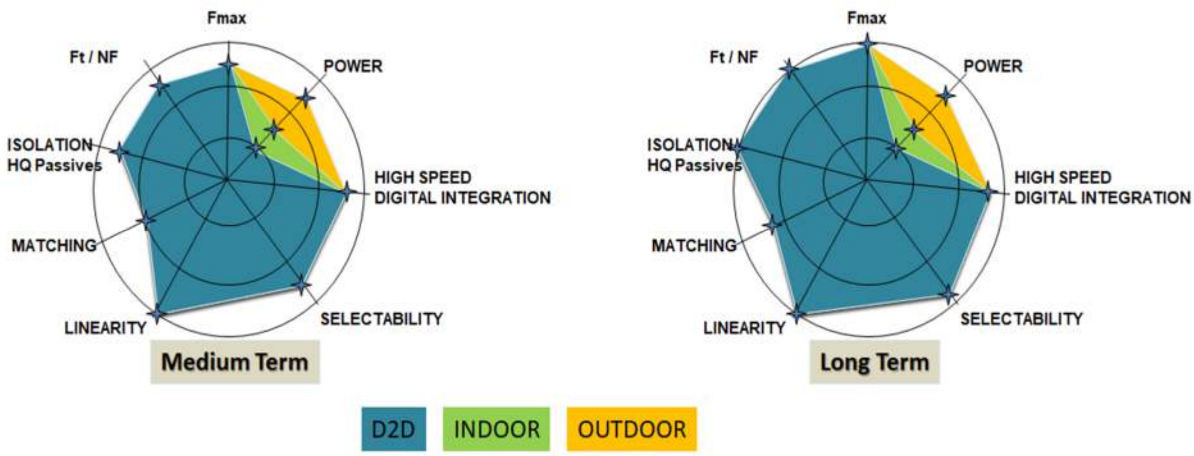


Figure 2.2.3 - Medium-term (2025) and long-term (2030) solid state technology roadmap proposed by H2020 CSA project NEREID[1]

Note that the vision presented in Figure 2.2.3 also applies to packaging and printed circuit board (PCB) technologies. It is also worth noting that while Europe is playing a key role in innovative differentiated semiconductor technologies, there is very little R&D activity or few players in Europe on the packaging and PCB side. This point is clearly a weakness that should be addressed to strengthen Europe's connectivity technology portfolio.

Securing Europe's differentiated hardware technology manufacturing

Beyond the development and enablement in Europe of innovative semiconductor technologies targeting the connectivity market, it will be key to safeguard and promote European manufacturing capability to both secure Europe economical interest (in terms of market share) and also address the sovereignty topic (since trade war issues can jeopardize the viability of Europe's industrial actors). To do so, in coordination with chapter 1.1 on Process Technology, Equipment, Materials and Manufacturing, the following topics should be supported:

- The enablement of pilot lines: the objective here is to support the deployment of additional manufacturing capabilities for technology already available in Europe (supporting the transition to 300 mm Fab), or to address new technologies (such as packaging or advanced PCB) to increase the technology portfolio available in Europe.
- The rise of new semiconductor equipment champions: to secure manufacturing capabilities in the long term, it will also be necessary to ensure that the required equipment is provided by European players. This is crucial to prevent any vulnerability in the European supply chain to possible international political or economic issues.

[1] https://www.nereid-h2020.eu/

- Nurture a pan-European design ecosystem to tackle the new challenges on transistors in more than Moore, circuits architectures with increased security and trustworthiness and new AI/ML chips raised by the digitization era.

Strengthening Europe's connectivity technology portfolio (hardware, internet protocols and software)

Leveraging previously discussed differentiated semiconductor technology portfolio, innovative connectivity solutions (hardware, internet protocol software) should be encouraged to enable Europe to take full advantage of its technology and manufacturing assets, and to capture market share at the component level. This action is crucial to secure Europe's position beyond 5G and 6G standardization and preliminary deployment activities. It also enables the development and manufacturing in Europe of highly integrated connectivity module/systems. Since most of the value of a complex connectivity system will be captured at the module level, it is highly desirable to enable European players to rise in the value chain (in coordination with the chapter Components, Modules and Systems Integration).

In targeting systems and applications, it is important to consider the interconnection between subsystems, and focus should be on individual component technology development according to needs identified at the system or application level. To support this system vision, the promotion of innovative technology enabling heterogeneous and chiplet integration is key.

Heterogeneous and chiplet integration refers to the integration of separately manufactured components into a higher- level assembly that cumulatively provides enhanced functionality and improved operating characteristics. In this definition, components should be taken to mean any unit – whether individual die, device, component, and assembly or subsystem – that is integrated into a single system. The operating characteristics should also be taken in their broadest meaning, including characteristics such as system-level cost of ownership.

This is especially true for the hardware side in the context of the end of Moore's law. It is the interconnection of the transistors and other components in the integrated circuit (IC), package or PCB and at the system and global network level where future limitations in terms of performance, power, latency and cost reside. Overcoming these limitations will require the heterogeneous integration of different materials (silicon, III-V, SiC, etc.), devices (logic, memory, sensors, RF, analogue, etc) and technologies (electronics, photonics, MEMS and sensors).

2.2.3.1.3 Key focus areas

To support the vision presented in the previous paragraph, we propose to focus effort on the following key focus areas:

- Materials (GaN, InP, etc.) and large diameter wafers technology (POI, InP on Si, GaN on Si, InP etc.) supporting the development of innovative connectivity technology solution.
- Differentiated semiconductor technology development targeting connectivity application.
- Packaging and PCB technology targeting connectivity application.
- Pilot line enablement to support the strengthening of European manufacturing capability.
- Connectivity solution engineering through virtualization of the different connectivity layers.
- Connectivity for high speed low energy Heterogeneous and chiplet integration
- Ultra-low power transceivers with low eco-footprint.
- Power efficient and cost efficient transceivers including Data conversion (ADC-DAC), up & down frequency conversion (LO & mixers), and RF emission and Reception (PA-LNA).
- Advanced System on Chip design for CMOS and new technologies like e.g. GaN, InP, InP on Si, GaN on Si.
- Innovative EDA solutions leveraging AI capability.
- The enablement of a EU fabless ecosystem which could leverage advanced semiconductor manufacturing in Europe and enable to strengthen EU position in the value chain.

2.2.3.2 Major Challenge 2: Investigate innovative wireless connectivity technology (new spectrum or medium) and new approaches to improving existing connectivity technology to maintain the EU's long-term leadership

2.2.3.2.1 State of the art

Targeting connectivity solutions beyond 5G, R&D activity is today mainly focused on the three key challenges listed below.

Evaluating the advantage to use new spectrum (especially 6 GHz – 15 GHz band and mmW frequencies >100 GHz)

With the ongoing deployment of 5G in the <6 GHz, the current R&D focus is now focused on the spectrum considered for 6G development targeting the 2030 time horizon. Three main frequency bands are today discussed in the telecommunication industry (following the recent decision of WRC23 in December 2023):

<u>Low band (6.425 GHz – 7.125 GHz):</u>
One of the significant outcomes of WRC-23 was the allocation of the 6.425 GHz – 7.125 GHz band as a licensed one dedicated to 6G while the lower 6 GHz spectrum (5.925 GHz – 6.425 GHz) will remain available for unlicensed usage (namely Wi-Fi). WRC-23 thus gave a balanced decision which supported both services. We can note that while this new spectrum will be the main 6G one identified by most region to start deployment in the 2030 time frame, the USA has adopted a complete opposite position. While China is considering to even allocate the lower 6 GHz spectrum for IMT services, the US did the opposite and the FCC allocated the full 5.925 GHz – 7.125 GHz band for Wi-Fi (raising the question of which low frequency spectrum may be used to deploy 6G in the USA). Development of wireless system in the 6.425 GHz – 7.125 GHz will likely be incremental to previous one <6 GHz, however we can mention that PA efficiency (GaN opportunity on the handset market?) and filtering technology (due to coexistence issues with Wi-Fi) seem to be the main challenges today.

<u>Mid band (7.125 GHz – 8.4 GHz):</u>
WRC-23 also set the agenda for the next World Radiocommunication Conference in 2027 (WRC-27) with a clear roadmap for future IMT spectrum allocations. One of the key identified new spectrum to be evaluate is the 7.125 GHz - 8.4 GHz one since it seems to be the most likely solution to add new spectrum for 6G deployment in the US (a federal spectrum study working group started in march 2024 with the goal of releasing a final report on potential commercial access to the band by October 2026).
Since it may prove difficult to clear this entire spectrum from existing usage, we can note that the 7.75 GHz – 8.4 GHz band is attracting a lot of attention since such scenario poses less technical difficulty and has the advantage to be considered by other region than the USA.
From hardware technology point of view, working over 8 GHz will bring several challenges. From PA side, we can wonder if GaAs HBT technology will remain relevant. From filter side, It is not clear if existing filter technologies will provide acceptable performances. More generally, it creates an opportunity for SiGe technology (and consequently for Europe) since this technology is today widely used in the X and Ku bands for satellite communication.

<u>NTN in Ku & Ka bands:</u>
In addition to new spectrum requirements previously mentioned, 6G envision to seamlessly integrate existing satellite communication in cellular connectivity networks. This trend can be seen under the topic Non Terrestrial Networks currently discussed inside 3GPP.
Targeted frequency spectrum is not new (traditional Ku and Ka bands). The main objective is here to leverage under deployment Low Earth Orbit satellite constellation to complement the cellular network coverage while offering performances (latency & data rate) in line with real-time application.
Two main challenges will have to be addressed: seamless integration and handover of satellite connectivity along with cellular one (network type should be transparent for the user) and the development of cost-effective chipset solutions to enable user terminals in line with mass market constraints (which creates key opportunities for Si-based technology such as SiGe BiCMOS).
We can note that on this NTN topic, Supplemental Connectivity from Space is a new trend to be kept in mind. The idea consists in reusing frequency bands used by cellular operators on the ground to complement their

existing cellular network with a LEO satellite connectivity. The main ongoing development concerns what SpaceX and T-Mobile are doing in the US. The advantage is here to leverage an installed base of user's equipment (4G smartphones) since the only new hardware is from the infrastructure side (the LEO satellite constellation). This development is today limited to what regulator are authorizing on their market (for example the FCC in the US), but if successful this approach could disrupt the current 3GPP view on NTN.

While the R&D development are today focused on frequencies below 15 GHz, there is some interest in assessing achievable performances with a higher frequency (even if now standardization activity are for the moment ongoing either from 3GPP or WRC side). D band (~140 GHz) remains a strong topic of interest. For regulatory reasons, the 275 GHz – 325 GHz range holds promise as it enables the widest available bandwidth. As an illustration, the US has facilitated their research on the 95 GHz – 3 THz spectrum over the coming decade. After a unanimous vote, the Federal Communications Commission (FCC) has opened up the "terahertz wave" spectrum for experimental purposes, creating legal ways for companies to test and sell post-5G wireless equipment. However, we can note that the telecommunication industry is cautious on the use of this spectrum. Magnus Frodigh, Ericsson's chief researcher, for example wrote[2]: "We have thus identified new potential spectrum ranges for 6G, notably in the centimetric range from 7-15 GHz, which we believe will be an essential range, and in the sub-THz range from 92-300 GHz, which will have a complementary role serving niche scenarios". He continued: "Our learnings from 5G are that the mmWave range is a powerful spectrum range which allows operators to provide value to industry and enterprise through high data rates. However, due to limited coverage, it serves as a complement to other ranges that can be used in wider areas but with limited data rates (i.e. mid bands)." Frodigh concluded: "We believe that in order to benefit society, the majority of 6G use cases should be enabled for wide-area coverage, both indoors and outdoors, and not limited to confined areas. This means that – whilst we ought to explore the sub-THz region for entirely novel 6G capabilities – the main value will be in the centimetric 7-15 GHz". At such high frequency, new hardware technology may be required and InP is today a hot topic to enable low noise receivers and power efficient Power Amplifiers > 100 GHz.

Exploring the benefits that AI could bring to connectivity technologies

While 5G is being deployed around the world, efforts by both industry and academia have started to investigate beyond 5G to conceptualize 6G. 6G is expected to undergo an unprecedented transformation that will make it substantially different from the previous generations of wireless cellular systems. 6G may go beyond mobile internet and will be required to support ubiquitous AI services from the core to the end devices of the network. Meanwhile, AI will play a critical role in designing and optimizing 6G architectures, protocols and operations.

For example, two key 5G technologies are software-defined networking (SDN) and network functions virtualization (NFV), which have moved modern communications networks towards software-based virtual networks. As 6G networks are expected to be more complex and heterogeneous, advanced softwarisation solutions are needed for beyond 5G networks and 6G networks. Selecting the most suitable computational and network resources and the appropriate dynamic placement of network functions, taking network and application performance as well as power consumption and security requirements into account, will be an important topic. By enabling fast learning and adaptation, AI-based methods will render networks a lot more versatile in 6G systems. The design of the 6G architecture should follow an "AI-native" approach that will allow the network to be smart, agile, and able to learn and adapt itself according to changing network dynamics.

Given the integration of wireless and wired and legacy industrial edge connectivity, heterogeneous network virtualisation is of increasing importance to the System of Systems integration in all application domains. Manageable virtual heterogeneous networks from edge to cloud will become a strong competitive advancement in the near future. Given the foreseen network complexity AI assisted and partly autonomous management is clearly important.

2.2.3.2.2 Vision and expected outcome

To address identified connectivity technology challenges, we propose the vision described below, which can be summarized in the following three points (with associated expected outcomes).

[2] https://www.ericsson.com/en/blog/2022/11/why-its-time-to-talk-6g

Assess achievable connectivity performances using new spectrum

To maintain European leadership on connectivity technology and ensure sovereignty, the development of new electronics systems targeting connectivity applications in non-already standardized (or in the process of being standardized) spectrum should be supported. A special focus should be dedicated to the frequency bands listed below:

- ***6G connectivity enablement in the 6.425 GHz – 7.125 GHz band***: This band having been identified as the main one to be deployed at short terms for 6G, it is of strategic importance to ensure both sovereign European hardware development and a timely spectrum allocation to operator to secure the position of EU in the 6G race.

- ***Investigation of additional 6G spectrum in the 7.75 GHz – 8.4 GHz band***: This band being the most likely one at mid terms to support 6G deployment in the US, and also being a serious option to offer additional 6G spectrum worldwide, EU has to ensure a pragmatic evaluation of this spectrum to both be able to play a role in WRC body standardization activities and secure the market share of EU wireless infrastructure players on the US wireless market.

- ***Investigation of Supplemental Connectivity from Space in existing 4G & 5G bands***: While NTN in Ku and Ka band is today a hot topic inside 3GPP, it can not be excluded that SCS approach pushed by SpaceX and T-Mobile may disrupt current NTN view (in a very pragmatic way). Consequently, SCS scenario evaluation in EU leveraging existing cellular 4G & 5G spectrum should be encouraged and dedicated R&D activities supported to secure this technical option in EU and ensure both a EU sovereign connectivity solution and possible business opportunities for EU actors.

- ***Investigate possible deployment of Wi-Fi-8 at 60 GHz***: In the context of Wi-Fi-8 standard definition, Intel and other players have been pushing the support of the 60 GHz band to enable mmW based Wi-Fi connectivity (and achieved higher data rate than Wi-Fi-7 in a power efficient manner). The IEEE 802.11 Integrated Millimeter Wave Study Group (IMMW SG) has been specifically created in 2023 to investigate this scenario and decide about the addition of the 60 GHz band to Wi-Fi standard. R&D activity on this topic in EU should be encouraged to ensure that if selected this license free frequency band could be leveraged in EU to enable affordable high speed connectivity to the market.

- ***Sub-THz connectivity application in D band (~140 GHz) and > 200 GHz***: With THz communication being a hot topic in the international academic community (especially in the US and China), European activity in the spectrum > 200 GHz should be encouraged (with a special focus on high speed point to point wireless link). These investigations should help Europe play a role in the development of the new technology and assess its relevance for future standardization activities.

Integrate AI features to make connectivity technology faster, smarter and more power-efficient

The use of new spectrum or propagation mediums is not the only way to boost innovative connectivity technology. As mentioned, 5G has underlined the role of software to promote virtualization and reconfigurability, but those concepts may not be sufficient to address the challenges related to the more complex connectivity technology that may be developed (for example, 6G).

To address this challenge, Artificial Intelligence is now perceived as a strong enabler. Consequently, in coordination with the "Edge Computing and Embedded Artificial Intelligence" chapter, the topics below should be supported:

- Investigate AI features at the edge: to improve the power efficiency of mobile devices by reducing the amount of data to be transmitted via the wireless network, the concept of AI at the edge (or edge AI) has been proposed. The idea is to locally process the data provided by the sensor using mobile device computing capability. Moreover, processing data locally avoids the problem of streaming and storing a lot of data to the cloud, which could create some vulnerabilities from a data privacy perspective.

- Use AI to make the connectivity network more agile and efficient: the idea here is to move to an AI-empowered connectivity network to go beyond the concept of virtualization and achieve new improvements in terms of efficiency and adaptability. For example, AI could play a critical role in designing and optimizing 6G architectures, protocols and operations (e.g. resource management, power consumption, improved network performance etc.).

2.2.3.2.3 Key focus areas

To support the vision presented in the previous paragraph, we propose that efforts should be focused on the following key focus areas:

- Innovative connectivity system design using new spectrum.
- Investigation and standardization activity targeting 6G cellular application in the frequency band < 15 GHz.
- Development of innovative connectivity technology using unlicensed frequencies in the 6 GHz – 15 GHz band.
- Development of connectivity system leveraging the concept of edge AI.
- Evaluation of the AI concept to handle the complexity of future connectivity networks (for example, 6G), and to improve efficiency and adaptability.

2.2.3.3 Major Challenge 3: Autonomous interoperability translation for communication protocol, data encoding, compression, security and information semantics

2.2.3.3.1 State of the art

Europe has a very clear technology lead in automation and digitalisation technology for industrial use. The next generation of automation technology is now being pushed by Industry 4.0/5.0 initiatives backed by the EC and most EU countries. In the automotive sector, the software defined vehicle initiative supports the autonomous and green car vision is the driver. Here, Europe again has a strong competitive position. In healthcare, the ageing population is the driver. Europe’s position in this area is respectable but fragmented. Robust, dependable, secure and interoperable connectivity from application to application and prepared for interaction in System of Systems solution are fundamental to market success in these and other areas.
Interoperability is a growing concern among numerous industrial players. An example here is the formation of industrial alliances and associated interoperability project efforts. One of the directions chosen targets is to gather behind a few large standards. An example of this is showcased in Figure 2.2.4.

To maintain and strengthen the European lead, advances in autonomous interoperability and associated efficient engineering capability are necessary. The game changers are:

- Autonomous interoperability for SoS integration for efficient machine supported engineering at design-time and run-time.
- Open interoperability infrastructure, and engineering support.
- Novel, flexible and manageable security solutions.
- Standardisation of the above technologies.

Figure 2.2.4 - ISO 15926 – Asset Standards worldwide (Source: Erik Molin, SEIIA)

2.2.3.3.2 Vision and expected outcome

To fully leverage heterogeneous integration at the hardware level, software interoperability is a parallel challenge to provide application to application connectivity that allows for autonomous SoS connectivity, from edge to cloud, enabling usage of available data for all areas of application. To do so, dedicated software tools, reference architecture infrastructure and standardisation are key to supporting autonomous interoperability, thus enabling the provision of a widely interoperable, secure, scalable, smart and evolvable SoS connectivity.

This challenge involves the interoperability of service or agent protocols, including encoding, security and data semantics. Here, payload semantics interoperability is a primary focus, leading to architectures, ontologies, technologies and engineering tools that support application to application integration of SoS for all areas of applications at design- time, in run-time and over life cycle. This will include e.g. integration of and translation between different standards used in domains where SoS interaction is necessary to reach business and societal objectives.

The objective here is a technology that enables nearly lossless interoperability across protocols, encodings and semantics, while providing technology and engineering support foundations for the low-cost integration of very large, complex and evolvable SoS.

Expected achievements are:

- Open source SoS architecture infrastructure with reference implementation supporting interoperability, security scalability, smartness and evolvability across multiple technology platforms, including 5G & 6G.
- Open source engineering and implementation infrastructure for the de-facto standard SoS connectivity architecture.
- Architecture reference implementations with performance that meets critical performance requirements in focused application areas.

2.2.3.3.3 Key focus areas

The high-priority technical and scientific challenges in both design-time and run-time are:

- Semantics interoperability from application to application.
- Autonomous translation of protocols, encodings, security and semantics.
- Evolvable SoS connectivity architecture infrastructure and technologies over time and technology generations.

2.2.3.4 Major Challenge 4: Architectures and reference implementations of interoperable, secure, scalable, smart and evolvable IoT and SoS connectivity

2.2.3.4.1 State of the art

It is clear that the US is the leader when it comes to wired IP-connectivity while Europe is the leader in cellular IP-connectivity. The big potential game changer here is 5G and upcoming 6G. To advance the European position, the establishment of connectivity architecture, reference implementation and associated engineering frameworks supporting 5G/6G and other wireless technologies is required. The integration of IP-connectivity with the vast industrial legacy edge connectivity needs a special focus. Such focus will foster a faster industrial and societal digitalisation enabling the integration of efficient and optimisable flexible production value networks. Given the longevity of such IP and legacy connectivity the life cycle perspective over technology life time and generations will need special considerations. In certain domains such as automotive and industrial automation, Europe is the major player. Market studies[3] indicate very large to extreme growth in the SoS market over the next five to ten years. This will provide a very strong market pull for all technologies and products upstream. Here, connectivity interoperability is a very important component, enabling tailored SoS solutions and efficient engineering. The vision is to provide interoperable connectivity architecture, reference implementation and associated engineering support and frameworks spanning technologies from legacy to 5G and 6G and other wireless and wired technologies.

2.2.3.4.2 Vision and expected outcome

The enabling of SoS interoperable connectivity is fundamental for capturing the emerging SoS market and its very high growth rate. Efficient engineering and the deployment of interoperable, secure, scalable, smart and evolvable SoS connectivity will be key to this. This will help Europe lead in the establishment of interoperable connectivity architecture, reference implementation and associated engineering frameworks.

The identified game changers are:

- Establishment of connectivity architecture infrastructure with associated reference implementation and related engineering frameworks and associated standardisation
- SoS application to application connectivity being interoperable, secure, scalable, smart and evolvable over technology generations.

Expected achievements

- Open source implementation of reference architecture infrastructure supporting interoperability, security scalability, smartness and evolvability across multiple technology platforms, like e.g. 5G/6G, wired and optical connectivity.
- Open source engineering and implementation frameworks for the de-facto standard SoS connectivity architecture infrastructure.
- Architecture infrastructure reference implementations which meet critical performance requirements in focused application areas.

[3] Advancy, 2019: Embedded Intelligence: Trends and Challenges, A study by Advancy, commissioned by ARTEMIS Industry Association. March 2019.

2.2.3.4.3 *Key focus areas*

The high-priority technical and scientific challenges are:

- SoS interoperable connectivity architecture infrastructure as a de-facto standard.
- Reference implementation of de-facto SoS connectivity architectures.
- Engineering frameworks for de-facto standard SoS connectivity architecture.

2.2.3.5 Major Challenge 5: Network virtualisation enabling run-time engineering, deployment and management of edge to cloud network architectures.

2.2.3.5.1 *State of the art*

Virtualisation of networks is a main trend for cellular networks. This has to be expanded to other wireless and wired connectivity technology including the vast industrial legacy edge connectivity. Current state of the art has a large technology differences between the legacy and IP based connectivity. This includes as well technology and operational competences in the industry and society. Thus wide and flexible digitalisation of industrial and societal value networks is rather limited.

2.2.3.5.2 *Vision and expected outcome*

The enabling of virtualised networks is fundamental for capturing the emerging SoS market and its very high growth rate. Efficient engineering, deployment and management of connectivity is a key enabler for interoperable, secure, scalable, smart and evolvable SoS. This will help Europe lead in the establishment of connectivity architecture, reference implementation and associated engineering frameworks.

European leadership in certain domains, mentioned above, will provide a very strong market pull for all technologies and products upstream. Here, virtualised connectivity is a very important component, enabling dynamic updates and rearrangements of SoS solutions. The vision is to provide virtualised connectivity across physical and mac layers spanning technologies from legacy to 5G and upcoming 6G.

The identified game changers are:

- Technologies for network virtualisation across multiple hardware and software layers using heterogeneous devices.
- Engineering, integration and management tools and methodologies for engineering and operation of virtualized legacy, IoT and SoS networks and service components of technology and system life cycle.
- Intelligent and run-time re-organisation of hardware platforms.

Expected achievements:

- Open source implementation of reference architectures supporting virtualised connectivity across multiple technology platforms from industry legacy to IP based 5G, 6G, wired and optical.
- Open source engineering and management frameworks for virtualised connectivity across multiple technology platforms from industry legacy to IP based 5G, 6G, wired and optical.
- Reference implementations with performance that meets critical performance requirements in focused application areas taking into consideration energy efficiency.
- Run-time re-organisation of hardware and software platforms.

2.2.3.5.3 *Key focus areas*

The high-priority technical and scientific challenges are:

- Virtual connectivity architecture supporting multiple technology platforms.
- Reference implementation of virtual connectivity architecture enabling very efficient application-level usage.
- Engineering, integration and management frameworks with tools for virtual connectivity architectures.

2.2.4 TIMELINE

The timeline for addressing the Major Challenges in this section is provided in the following table.

MAJOR CHALLENGE	TOPIC	SHORT TERM (2025–2029)	MEDIUM TERM (2030–2034)	LONG TERM (2035 and beyond)
Major Challenge 1: Strengthening the EU connectivity technology portfolio to maintain leadership, secure sovereignty and offer an independent supply chain	Topic 1.1: Materials (GaN, InP, etc.) and large diameter wafers technology (POI, InP on Si, GaN on Si, InP etc.) supporting the development of innovative connectivity technology solution.	TRL 3–4 Enable innovative GaN (on Si or on SiC), InP (on Si or on bulk InP) and POI technologies	TRL 5–6 Industrial transfer of previous technologies from pilot line to Fab	
	Topic 1.2: Differentiated semiconductor technology development targeting connectivity application.	TRL 3–4 Enable next generation RF SOI and BiCMOS technology	TRL 5–6 Industrial transfer of previous technologies from pilot line to Fab	TRL 7–9 Support of high-volume production product in Europe
	Topic 1.3: Packaging and PCB technology targeting connectivity application.	TRL 4–6 Development of innovative European packaging (such as AMP or FOWLP) and PCB technologies (targeting HDI)	TRL 7–9 Pilot line enablement and support of small series	
	Topic 1.4: Connectivity solution engineering through virtualisation of the different connectivity layers	TRL 3–4 Proof of concept using existing COST hardware	TRL 5–6 Dissemination through standardization bodies	TRL 7–9
	Topic 1.5: Connectivity for high speed low energy Heterogeneous and chiplet integration	TRL 3–4 Proof of concept at elementary function level (LNA, PA, VCO, mixer, …) leveraging pilot line available in Europe	TRL 5–6 Higher complexity demo at the level of a full transceiver	TRL 7–9

	Topic 1.6: Ultra-low power transceivers with low eco-footprint	TRL 4–6 Proof of concept at full transceiver level	TRL 7–9 Transfer to the industry	
Major Challenge 2: Investigate innovative connectivity technology (new spectrum or medium) and new approaches to improving existing connectivity technology to maintain the EU's long-term leadership	Topic 2.1: innovative connectivity system design using new spectrums	TRL 3–4 Assess the specification of wireless systems in new 6G band (< 7 GHz and prospective mmW band > 100 GHz)	TRL 5–6 Achieve preliminary transceiver demonstrator using European technologies	TRL 7–9 Transfer to the industry to enable European products
	Topic 2.2: investigation and standardisation activity targeting 6G cellular application in frequency band < 10 GHz	TRL 3–4 Contribute to 6G new spectrum standardization (especially the 7.1 GHz – 8.4 GHz band targeted by WRC27)	TRL 5–6 Leverage European differentiated technologies to address 6G mid band spectrum challenges	TRL 7–9 Enable first 6G connectivity chipset solution
	Topic 2.3: development of innovative connectivity technology using unlicensed frequency in the 6 GHz – 7 GHz and 60 GHz band	TRL 4–6 Contribute to Wi-Fi8 standardization (especially on low power and Integrated Millimeter Wave features)	TRL 7–9 Enablement of European Wi-Fi 7 and 8 chipset solution leveraging European derivative technology	
	Topic 2.4: development of connectivity systems leveraging the concept of edge AI	TRL 3–4 Proof of concept using existing COST hardware	TRL 5–6 Dissemination through standardization bodies	
	Topic 2.5: evaluation of the AI concept to be able to handle the complexity of future connectivity networks (for example, 6G), and to improve efficiency and adaptability	TRL 3–4 Proof of concept using existing COST hardware	TRL 5–6 Dissemination through standardization bodies	
Major Challenge 3: Autonomous interoperability translation for communication protocol, data encoding, compression, security and information	Topic 3.1: semantics interoperability from application to application.	AI-supported translation of payload semantics based on a limited set of ontologies and semantics standards enabling application information usage	Scalability to payload semantics integration and translation across domain standards	Scalability to payload semantics integration and translation between application domains

semantics	Topic 3.2: autonomous translation of protocols, encodings, security and semantics	Autonomous and dynamic translation between SOA-based services protocol, data encoding, data compression and data encryption	Dynamic translation between major data model relevant for the ECS supported field of applications.	Autonomous and dynamic translation between a large set of data models relevant for the ECS supported field of applications
	Topic 3.3: evolvable SoS connectivity architecture infrastructure and technologies over time and technology generations	AI tool supported evolution of connectivity infrastructure and technologies over time and technology generations	AI and formal tool supporting evolution of connectivity infrastructure and technologies over time and technology generations	AI and formal tool supporting V&V of evolution of connectivity infrastructure and technologies over time and technology generations
Major Challenge 4: architectures and reference implementations of interoperable, secure, scalable, smart and evolvable IoT and SoS connectivity	Topic 4.1: SoS interoperable connectivity architecture infrastructure as a de-facto standard	SoS interoperable connectivity architecture based on SOA established as a major industrial choice within application domains of the ECS-SRIA	SoS interoperable connectivity architecture based on SOA established as the major industrial choice in-between the application domains of the SRIA	SoS interoperable connectivity architecture based on SOA established as the major industrial and societal choice
	Topic 4.2: reference implementation of de facto SoS connectivity architectures	Reference implementation of the SoS connectivity architecture becoming a natural part of the global SoS architecture (chapter SoS) reference implementation	Reference implementation of the SoS connectivity architecture becoming a natural part of the global SoS architecture (chapter SoS) reference implementation at TRL 8–9.	Reference implementation of the SoS connectivity architecture including semantic data model interoperability becoming a natural part of the global SoS architecture (chapter SoS) reference implementation
	Topic 4.3: engineering frameworks for de facto standard SoS connectivity architecture	Reference implementation of an engineering framework with associated tools for SoS interoperable connectivity	Standardised reference implementation of an engineering framework with associated tools for SoS interoperable connectivity	Standardised reference implementation of an engineering framework with associated tools for SoS interoperable connectivity including wide data model translation.

Major Challenge 5: network virtualisation enabling run-time engineering, integration, deployment and management of edge and cloud network architectures	Topic 5.1 Virtual connectivity architecture supporting multiple technology platforms,	A fully distributed edge to cloud virtual connectivity environment with a limited set of technology platforms	A fully distributed edge to cloud virtual connectivity management environment including legacy and IP based technology platforms.	A fully distributed edge to cloud virtual connectivity management environment across legacy and IP based technology platforms.
	Topic 5.2 Reference implementation of virtual connectivity architecture enabling very efficient application-level usage.	Edge to cloud , integration of limited number of connectivity technology platforms supporting virtualisation and programmability	Pilot scale demonstrations of fully distributed edge to cloud environments, including hardware accelerators to significant lower cost levels compared to current industrial SOTA	Pilot scale demonstrations of fully distributed edge to cloud environments covering from legacy to IP based connectivity
	Topic 5.3 Engineering, integration and management frameworks	Virtualisation engineering, integration and management tools and tool chains for Edge to cloud , integration of limited number of connectivity technology	Virtualisation engineering, integration and management tools and tool chains for Edge to cloud pilot scale integration of limited number of connectivity technology	Virtualisation engineering, integration and management tools and tool chains for Edge to cloud environments covering from legacy to IP based connectivity

2.3

Cross-Sectional Technologies

ARCHITECTURE AND DESIGN: METHODS AND TOOLS

2.3 Architecture and Design: Methods And Tools

2.3.1 Scope

Two assets are essential to strengthen European industry's potential to transform new concepts and ideas cost- and effort-effectively into high- value and high-quality innovations and applications based on electronic components and systems (ECS): Effective architectures and platforms at all levels of the design hierarchy; and structured and well-adapted design methods and development approaches supported by efficient engineering tools, design libraries and frameworks. These assets are key enablers to produce ECS-based innovations that are: (i) beneficial for society; (ii) accepted and trusted by end-users; and thus (iii) successful in the market. Next to these technologically induced advancements and benefits, the methods in this chapter also further the European goal of supporting sustainability (c.f. Chapter 0), both by advancing the creation of sustainable components and systems as well as supporting their creation in a sustainable way.

Future ECS-based systems will be intelligent (using intelligence embedded in components), highly automated up to fully autonomous, and evolvable (meaning their implementation and behaviour will change over their lifetime), cf. Part 3. Such systems will be connected to, and communicate with, each other and the cloud, often as part of an integration platform or a system-of-system (SoS, cf. chapter 1.4). Their functionality will largely be realised in software (cf. chapter 1.3) running on high-performance specialised or general-purpose hardware modules and components (cf. chapter 1.2), utilising novel semiconductor devices and technologies (cf. chapter 1.1). This Chapter describes needed innovations, advancements and extensions in architectures, design processes and methods, and in corresponding tools and frameworks, that are enabling engineers to design and build such future ECS-based applications with the desired quality properties (i.e. safety, reliability, cybersecurity and trustworthiness, see also chapter 2.4, in which these quality requirements are handled from a design hierarchy point of view, whereas here a process oriented view is taken). The technologies presented here are therefore essential for creating innovations in all application domains (cf. Part 3); they cover all levels of the technology stack (cf. Part 1), and enable efficient usage of all cross-cutting technologies (cf. Part 2).

Due to the sheer size and complexity of current and future ECS-based products, the amount of functionality they perform, and the number and diversity of subsystems, modules and components they comprise, managing complexity and diversity have always been crucial when designing, implementing and testing these products. In addition, many of these systems need to fulfill high quality requirements, i.e. their performance, their usability, their dependability and in many cases also their functional safety and security need to conform to the highest standards. The trend of further growing functionality, complexity and diversity in future ECS-based applications combined with the advancement of new technologies (for example Artificial Intelligence) and system architectures (for example cloud- or edge-based architectures) further increases the corresponding challenges. Thus, engineers need to be able to create ECS-based products of increasing complexity fulfilling increasingly demanding challenges on the one hand, while still working cost- and effort-effective. In order to enable engineers to do so, design processes are in constant need to be adapted and to incorporate new methods and tools as well as completely new approaches to design and validation of ECS.

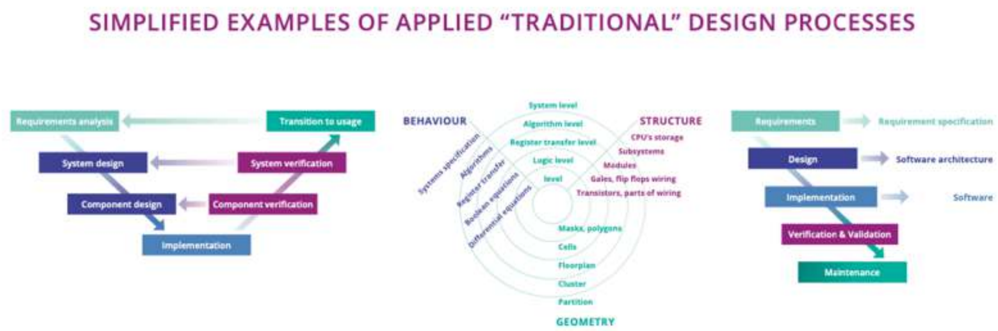


Figure 2.3.1 Simplified examples of applied "traditional" design processes: V-Model, Gajsky–Kuhn diagram (Y-chart) and the waterfall model. These are heavily in use, but not sufficient to handle future ECS-based systems and products.

One major change in design processes in recent years is, that the essentially linear traditional processes that typically end at market introduction (see Figure 1) are extended to so called continuous design processes that (a) span the whole lifecycle of a product and (b) use data collected from production, operation, and maintenance to improve on the original design. Data collected is used to (i) enable continuous updates and upgrades of products; (ii) enable in-the-field tests of properties that cannot be assessed at design-, development- or testing-time; and (iii) increase the effectiveness of validation and test steps by virtual validation methods based on this data (see also Major Challenge 2 and 3 in chapter 1.3 Embedded Software and Beyond). Apart from the technical challenges in collecting and analysing this data and/or using it for Maintenance purposes, non-technical challenges include compliance to the appropriate data protection regulations and privacy concerns of system's owners (Intellectual Property) and users (privacy data).

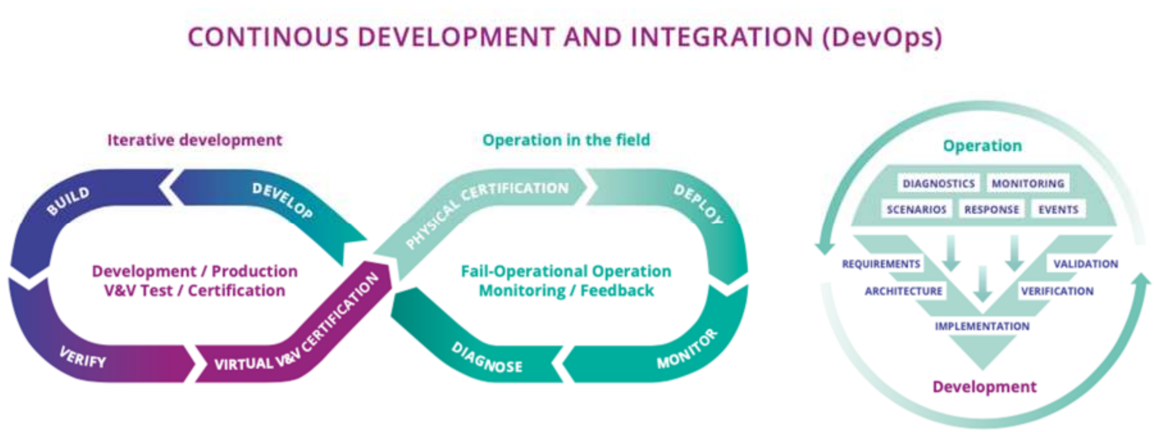


Figure 2.3.2 Simplified examples for continuous development processes (DevOps processes). Such processes are essential for building future ECS-based systems and products since they enable data collected during the operation phase to be used in iterative (continuous) development for updates of existing products.

Second, while model-based design techniques have already for some time been a major instrument to decrease both design complexity and corresponding validation effort, they now are a main enabler also for the shift to virtual engineering and virtual validation. This shift, in turn, is essential for cost- and effort-effective design and is also mirrored in current regulations for e.g. type certification. With virtual engineering, much of the design effort can be done 'virtually', i.e. by creating and using digital twins and formal models and/or simulation based test and validation methods. Additional supporting techniques are divide-and-conquer based approaches, both on a technical level – where modular, hierarchical designs need to be integrated into reference architectures and platforms –, and also on an organizational level – i.e. by employing open source solutions like e.g. RISC-V (cf. appendix A) or the various open-source integration platforms (cf. chapter 1.3 and 1.4), to increase interoperability and thus cooperation.

The resulting agile "continuous development processes" will ease quality properties assurance by providing design guidelines, design constraints and practical architectural patterns (e.g. for security, safety, testing), while giving engineers the flexibility and time to deliver the features that those development methodologies support ("quality-by-design solutions"). Last, but not least, the topic of virtual validation is of central importance in these continuous development processes, since both, the complexity of the system under test and the complexity and openness of the environment in which these systems are supposed to operate, are prohibitive for validation based solely upon physical tests. Although considerable advances have been made recently in scenario based testing approaches, including scenario generation, criticality measures, ODD (Operational Design Domain) definitions and coverage metrics, simulation platforms and testing methodologies, and various other topics, further significant research is needed to provide complete assurance cases as needed for certification, which combine evidences gained in virtual validation and verification with evidences generated in physical field testing to achieve the high confidence levels required for safety assurance of highly automated systems.

ECS-based applications are becoming increasingly ubiquitous, penetrating every aspect of our lives. At the same time, they provide greater functionality, more connectivity and more autonomy. Thus, our dependency on these systems is continuously growing. It is therefore vitally important that these systems are trustworthy – i.e. that they are guaranteed to possess various quality properties (cf. chapter 1.4) like safety, security, dependability, reliability and similar. Trustworthiness of ECS-based applications can only be achieved by implementing all of the following actions.

- Establishing architectures, methods and tools that enable "quality by design" approaches for future ECS-based systems (this is the objective of this chapter). This action comprises:
    - Providing structured design processes, comprising development, integration and test methods, covering the whole system lifecycle and involving agile methods, thus easing validation and enabling engineers to sustainably build these high-quality systems.
    - Implementing these processes and methods within engineering frameworks, consisting of interoperable and seamless toolchains providing engineers the means to handle the complexity and diversity of future ECS-based systems.
    - Providing reference architectures and platforms that ensure interoperability, enable European Industries to re-use existing solutions and, most importantly, integrate solutions from different vendors into platform economies.
- Providing methodology, modelling and tool support to ensure that all relevant quality aspects (e.g. safety, security, dependability) are designed to a high level (end-to-end trustworthiness). This also involves enabling balancing trade-offs with those quality aspects within ECS parts and

for the complete ECS, and ensuring their tool-supported verification and validation (V&V) at the ECS level.

- Providing methodology, modelling and tool support to enable assurance cases for quality aspects – especially safety – for AI-based systems, e.g. for systems in which some functionality is implemented using methods from Artificial Intelligence. Although various approaches to test and validate AI-based functionality are already in place, today these typically fall short of achieving the high level of confidence required for certification of ECS. Approaches to overcome this challenge include, amongst others:
    - Adding quality introspection interfaces to systems to enable engineers, authorities and end-users to inspect and understand why systems behave in a certain way in a certain situation (see "trustworthy and explainable AI" in chapters 2.1 and 2.4), thus making AI-based and/or highly complex system behaviour accessible for quality analysis to further increase user's trust in their correctness.
    - Adding quality introspection techniques to AI-based algorithms – i.e., to Deep-Neural Networks (DNN) – and/or on-line evaluation of 'distance metrics' of input data with respect to test data, to enable computation of confidence levels of the output of the AI algorithm.
    - Extending Systems Engineering methods – i.e., assurance case generation and argumentation lines – that leverage the added introspection techniques to establish an overall safety case for the system.

2.3.2 Major Challenges

We identified four **Major Challenges** within the transversal topic "Architecture and Design: Methods and Tools".

- **Major Challenge 1**: Enabling cost- and effort-efficient Design and Validation Frameworks for High Quality ECS. The ever-increasing functionality of ECS, usage and integration of new technologies to enable these functions and the high demands for validation and testing to ensure their quality drive the need for efficient, framework- and tool-supported design and validation processes and frameworks.
- **Major Challenge 2:** Enabling Sustainable Design for Sustainability. Methods and tools to support the design and validation of sustainable ECS as well as supporting a sustainable design and validation process.
- **Major Challenge 3**: Managing complexity. This challenge deals with methods to handle the ever-increasing complexity of ECS-based systems.
- **Major Challenge 4**: Managing diversity. Handling diversity in all aspects of developing ECS-based systems is the key objective of this challenge.

Together, these four challenges answer the need for Software tools and frameworks for engineering support for sustainable high-quality ECS covering the whole lifecycle.These four challenges are highly interconnected, and can hardly be seen separately: First, advancements in one challenge often also implies advancements in other challenges. For example, techniques to reduce the complexity of the ECS-based system from Major Challenge 3 often also result in a decrease of cost and effort for validation and test within the design and validation process (and corresponding framework), as described in Major Challenge 1. Second, solutions in one Major Challenge often also require at least partial solutions in other Major Challenges to be applicable.

2.3.2.1 Major Challenge 1: Enabling cost- and effort-efficient Design and Validation Frameworks for High-Quality ECS.

2.3.2.1.1 State of the art

Future ECS-based systems need to be connected, intelligent, highly automated, and even autonomous and evolvable. This implies a huge variety of challenges, including how to validate autonomous systems (considering that a full specification of the desired behaviour is inherently impossible), how to test them (critical situations are rare events, and the number of test cases implied by the open-world-assumption is prohibitively large), and how to ensure safety and other system quality properties (security, reliability, availability, trustworthiness, etc.) for updates and upgrades.

With the increased functionality of ECS and with the increasing demands on testing and validating them, engineers need tool support to enable them to design, build, test and validate, produce and maintain such complex system; these tools need to be integrated into seamless design flows (frameworks) such that the effort and the cost associated with these tasks ideally stays constant even though functionality and validation demands increase. New functionality of ECS, that cannot be handled by currently available tools, requires new methods which, when successful, lead to additional design and validation tools.

Both technology push that enables new functionalities and – even more – market pull, i.e. the demand for new functions in products – are extremely high in ECS. This is why design methods often lag behind current development methods. This holds even more for validation tools, where there a prominent examples – like showing safety for systems with functions based on AI – for which even the methods used are not complete, yet, let alone their implementation in appropriate tools. This leads to the absurd situation, that although the technologies needed for specific functions are available and even though the market demands these functions, Engineers cannot deliver them at all, or not in the required quality (i.e., safety) or not at an affordable price (because the effort, and thus the cost, for the manufacturer to build and validate these Systems is too high.

2.3.2.1.2 Vision and expected outcome

The vision is to enable European engineers to extend design processes and methods to a point where they allow handling of future ECS-based systems with all their new functionalities and capabilities over the whole lifecycle. Such extended processes must retain the qualities of the existing processes: (i) efficiency, in terms of effort and costs; (ii) enable the design of trustworthy systems, meaning systems that provably possess the desired quality properties of safety, security, dependability, etc.; and (iii) be transparent to engineers, who must be able to completely comprehend each process step to perform optimisations and error correction.

Such extended processes will cover the complete lifecycle of products, including production, Maintenance, decommissioning, and recycling, thereby allowing continuous upgrades and updates of future ECS- based systems that also address the sustainability and environmental challenges (i.e. contribute to the objectives of the Green Deal, see also Major Challenge 2). The main objectives to be reached here are the creation and extension of Lifecycle-aware holistic design flows that especially allow continuous development and updates of systems by collecting data from operations and maintenance (and possibly other phases of the process), feeding it back into the design phase and allowing for continuous improvements and updates. Equally important is the switch to virtual engineering of ECS, where many of the test and validation efforts required are done by simulation and analysis methods, thereby considerably increasing the efficiency of this. Advances in System Design Methods, especially in Model-Based Design, as well as in V&V methods complete the key research areas that are most important for this challenge.

2.3.2.1.3 Key focus areas

This Major Challenge comprises the following key focus areas:

Lifecycle-aware holistic design flows

"Closing the loop" – i.e. collecting relevant data in the operation phase, analysing it (using AI-based or other methods) and feeding it back into the development phase (using digital twins, for example) – is the focus of this research topic. It is closely related to the major challenges "Continuous integration and deployment" and "Lifecycle management" in chapter 1.3, which examines the software part of ECS, and Major Challenges 1, 2 and 4 in chapter 2.4.

Closing the loop includes data collected during operation of the system on all levels of the hierarchy, from new forms of misuse and cyber-attacks or previously unknown use cases and scenarios at the system level, to malfunctions or erroneous behaviour of individual components or modules. Analysing this data leads to design optimisations and development of updates, eliminating such errors or implementing extended functionality to cover "unknowns" and "incidents". On the one hand, this continuous development allows for shared learning, where in a system family (like a fleet of highly automated cars) all systems can be optimized based the experience of each single member of the family. Combined with a strong monitoring concept – i.e., where unknown scenarios and unknown use cases are detected before they actually arise and where corresponding safety measures like minimum risk maneuvers or evasive actions can be triggered – this 'learning in the field' concept can achieve significant advancements in safety assurance cases of such systems (online validation and verification).

All of these aspects must be supported within holistic design flows (frameworks) that also must support:

- Supply-chain-Awareness: From requirements to optimized system architecture considering supply chain leveraging seamless digital twin from component to design to manufacturing to operation

- Complete traceability of products and processes in virtual engineering, supporting sensitivity analysis and robustness investigation, included in the optimization process and the system monitoring process
- Design for optimized manufacturing and operation; awareness of physical effects and interferences; awareness of complete lifecycle, incl. energy, resource, CO2-footprint, recycling, circular economy
- Consistent methods and new approaches for (multi-level, multi-paradigm) modeling , analysis, verification and formalization of ECS's operational reliability and service life
- Open (and inner) source in HW and SW for complete product lifecycle

As non- (or partly-) technical Challenges, all data collection activities described in this chapter also need to comply to privacy regulations (e.g. the General Data Protection Regulations GDPR of the EU) as well as in a way that protects the Intellectual Property (IP) of the producers of the systems and their components.

Virtual engineering of ECS

Design processes for ECS must be expanded to enable virtual engineering on all hierarchy levels (i.e. from transistor level "deep down", up to complete systems and even System of Systems, cf. "Efficient engineering of embedded software" in chapter 1.3 and "SoS integration along the lifecycle" in chapter 1.4 for more details of this software-focused challenge, especially with respect to SoS). This requires model-based design methods including advanced modelling and specification capabilities, supported by corresponding modelling and specification tools. Furthermore, it is important to create reusable, validated and standardised models and model libraries for system behaviour, system environment, system structure with functional and non-functional properties, SoS configurations, communication and time-based behaviour, as well as for the human being (operator, user, participant) (cf. chapter 1.4 and the following key focus area "*Advancing System and component design (methods and tools))*.

Central to this approach are "digital twins", which capture all necessary behavioural, logical and physical properties of the system under design in a way that can be analysed and tested (i.e. by formal, AI-based or simulation based methods). This allows for optimisation and automatic synthesis (see also Major Challenge 1 and 2 in chapter 2.4, 'virtual prototypes' in appendix A, – for example, of AI- supported, data-driven methods to derive (model) digital twins.

Supporting methods include techniques to visualize V&V and test efforts (including their progress), as well as sensitivity analysis and robustness test methods for different parameters and configurations of the ECS under design. Test management within such virtual engineering processes must be extended to cover all layers of the design hierarchy, and be able to combine virtual (i.e. digital twin and simulation-based) and physical testing (for final integration tests, as well as for testing simulation accuracy).

To substantially reduce design effort and costs, a second set of supporting methods deals with the automated generation of design artefacts such as identification and synthesis of design models, automatic scenario, use-case and test vector generation, generative design techniques, design space exploration, etc. Typically, these build upon AI-supported analysis of field data.

Last, but not least, virtual validation and testing methods must be enhanced considerably in order to achieve a level of realism and accuracy (i.e., conformity to the physical world) that enables their use in safety assurance cases and thus fully enables the shift from physical testing to virtual testing. This includes overcoming limitations in realism of models and simulation accuracy, as caused for example by sensor phenomenology, vehicle imperfections like worn components, localization and unlimited diversity in traffic interactions.

Advancing System and component design (methods and tools)

System and component design, validation and test methods have to be continuously advanced and extended in order to keep up with the new functionalities of and new technologies used for ECS. Currently, advancement in the following topics is needed urgently to realize the potential that these new functionalities and technologies can bring:

- Model based design technologies
    - Model creation/elicitation, modelling techniques, modelling tools, model libraries, using explicit models as well as data-driven models
    - (AI-based) Model identification, synthesis, improvement and parameterization with measurement data
    - Techniques and tools to model behaviour, timing, functional and non-functional properties of (a) components, (b) systems, (c) environment / real world, (d) test-cases / scenarios (physical rule based as well as data-driven).
    - Executable models of sensors (incl. accuracy, confidence,…)
- Test-Management on all hierarchy layers
    - Automatic generation of test cases on all hierarchy layers (from physical sensor input via bitvectors up to concrete test scenarios for systems)
    - Means to efficiently process & analyse dynamic test results (traces, observations , loggings,…) to derive tangible knowledge for design improvements
- Augmented and virtual reality in design, development, manufacturing and maintenance processes
- Monitoring Techniques
    - V&V extended by life-time monitoring of security and reliability aspects
    - Methods and tools for (automatically generated) monitoring of systems (based on their digital twins), including monitoring for anomaly detection (for both security and safety)

In addition, there are two research topics requiring special attention:

First is the handling of uncertainty. The higher level of automatism – up to autonomy – of systems like cars and other transportation systems, medical machinery, production plants and many others, implies that the level of uncertainty that the system has to cope with also increases and has to be handled. This is mainly the case in perception, where the system ‘observes’ its surroundings with sensors. Despite long known and used techniques like sensor fusion and similar, there is always an inherent level of uncertainty as to the accuracy of the sensor data, which leads to ‘ghost artefacts’, i.e. detection of objects that are nor there in reality, and non-detection of objects that are. The other source of uncertainty is typically in prediction, where the system needs to predict how its surroundings will evolve over time in order to adjust its own behaviour accordingly. A key research topic here is therefore the advancement of design methods and V&V techniques for handling these uncertainties, where recent approaches implementing and guaranteeing bounds of these uncertainties and have shown first successes.

Second, the usage of Artificial Intelligence in Design and Validation promises a high potential for cost- and effort-efficiency, in a similar way as the use of AI in the systems themselves promises and increase in functionality of these systems.

- Integration of AI (including generative AI) and AI-based tools into engineering and development processes on all levels of the design hierarchy, to shorten development time, incl. metrics for quantification of covered design space, etc.
- Usage of AI and AI-based tools for V&V and development task.
- AI support for Model identification, synthesis, improvement and parameterization with measurement data (c.f. Model based design technologies above)
- Design and V&V for new technologies (i.e., flexible electronics, textile electronics…)

Usage of AI techniques in Design and Validation imposes the additional challenge that these have to be either supervised (i.e., they need to be able to explain to their user, why and how they achieved the results that they produce) or have to be validated themselves ('Who validates the validator?').

Integration of new V&V methods

The required changes of current design processes identified above, as well as the need to handle the new systems capabilities, also imply an extension of current V&V and test methods. First, safety cases for autonomous systems need to rely on an operational design domain (ODD) definition – i.e. characterisation of the use cases in which the system should be operated, as well as a set of scenarios (specific situations that the system might encounter during operation) against which the system has actually been tested. It is inherently impossible for an ODD to cover everything that might happen in the real world; similarly, it is extremely difficult to show that a set of scenarios cover an ODD completely. Autonomous systems must be able to detect during operation whether they are still working within their ODDs, and within scenarios equivalent to the tested ones. V&V methods have to be expanded to show correctness of this detection. Unknown or new scenarios must be reported by the system as part of the data collection needed for continuous development. The same reasoning holds for security V&V: attacks – regardless of whether they are successful or not – need to be detected, mitigated, and reported on. cf. chapter 1.4 and chapter 2.4)

Second, the need to update and upgrade future ECS-based systems implies the need to be able to validate and test those updates for systems that are already in the field. Again, corresponding safety cases have to rely on V&V methods that will be applied partly at design-time and partly at run-time, thereby including these techniques into continuous development processes and frameworks. For both of these challenges, energy- and resource-efficient test and monitoring procedures will be required to be implemented.

Third, V&V methods must be enhanced in order to cope with AI-based ECS (i.e., systems and components, in which part of the functionality is based upon Artificial Intelligence methods). This includes, amongst others, adding quality introspection techniques to AI-based algorithms – i.e., to Deep-Neural Networks (DNN) – and/or on-line evaluation of 'distance metrics' of input data with respect to test data, to enable computation of confidence levels of the output of the AI algorithm (compare to 'Explainable AI' in chapter 2.1) as well as extending Systems Engineering methods – i.e., assurance case generation and argumentation lines – that leverages the added introspection techniques to establish an overall safety case for the system.

2.3.2.2 Major Challenge 2: Enabling Sustainable Design for Sustainability

2.3.2.2.1 State of the art

Sustainability is a major goal of the European ECS industry; as described in Chapter 0, there are two sustainability goals:

On the one hand, we need to produce *sustainable ECS*, i.e. products that contribute to sustainability by being highly energy efficient, by producing less or even no emissions; products, which have a long lifetime and which are built from materials that can be recycled or used in other contexts afterwards (second life).

On the other hand, we need to *produce ECS in a sustainable way*, i.e., by using less energy for design, validation and production, and by increasing the useful lifetime of systems and/or their components.

Together, these actions contribute to eight of the seventeen sustainability goals of the United Nations and the EU, namely SDG 3 Good Health and Wellbeing, SDG 6 Clean Water (and sanitation), SDG 7 Affordable and clean Energy, SDG 8 Decent work and economic growth, SDG 9 Industry, Innovation, and Infrastructure, SDG 11 Sustainable Cities and communities, SDG 12 Responsible consumption and production, and SDG 13 Climate Action.

This Major Challenge groups the most prominent Key Focus Areas contributing to these goals. It should be noted, however, that these two goals have been a driving factor for many of the advancements done in ECS Design and Validation during the last decades. Building engines that are better – stronger, more reliable, etc. – while using less fuel or other resources, heavily contributing to the Smart Grid, usage of new materials and many other activities during recent years aim to achieve sustainable ECS based products as well as a sustainable way of designing and testing them. Thus, one kind find sustainability goals behind many of the focus areas described in Major Challenges 1, 3 and 4, although there sustainability might not be the main driver, but a highly advantageous second goal of that topic area.

2.3.2.2.2 Vision and expected outcome

There is a strong technology push combined with a need raised by the increased functionality of ECS based products to switch to lifecycle aware design processes and continuous development of ECS-based products, including data collection from systems under production and from operation, as described Major Challenge 1. This switch perfectly matches the Lifecycle aware Design Optimizations described in the following section. Together, these will enable the design of ECS based products that are produced using less and less energy and other resources, which are more easily repairable and which are recyclable or reusable at the end of their initial lifetime. Continuous Design Flows also drive the need for Updates and updatable systems, which perfectly matches the fact that Updates are also a need to extend useful lifetime of a product, thus increasing its sustainability. The two focus areas of 'Energy and resource efficient test procedures and equipment' and 'Low power design' directly contribute to the sustainability of the design and test process as well as the energy consumption of the system.

2.3.2.2.3 Key focus areas

Lifecycle aware Design Optimizations

Lifecycle aware design optimizations is a collective term for procedures and methods that aim at designing sustainable ECS. It comprises

- 'Design for producability' aiming at designing ECS that can be produced resource efficiently (i.e. with low consumption of energy and other resources'), Notice, that this topic comprises a lot more than only thinking about production materials and changing them in a sustainable way. For example, in the Electric-/Electronic Architecture of cars, the recent change from using point to point communication lines between different computational units to installing more and more communication busses (i.e. common communication lines used by many computational units) resulted not only in using less copper wires, but also in a fundamental shift in communication patterns and protocols that influenced the complete design of the application software.
- 'Design for recycling/reuse', which is mostly concerned with the materials, and
- 'Design for reparability', which comprises both, architectural methods that enable or ease reparability of a system as well as design and management of spare parts.

All of them essentially face the same challenges, namely

- Specification techniques for the corresponding requirements (of producibility, recycling/reuse and reparability)
- Requirement capturing
- Implementation techniques for these requirements
- Validation/Test methods for these requirements
- Conflict resolution techniques for requirement (i.e., how to handle a situation in which e.g., a safety requirement conflicts with a producibility requirement).

Especially for requirement capturing and validation methods, these techniques profit from data collection during production and operation, as described in Major Challenge 1 under "Lifecycle aware holistic design flows", which enables a continuous design flow also for these aspects of the system.

Updates

Updates resp. the property to be 'updateable' has the potential of significantly increasing the lifetime of a product and thus increase its sustainability. An 'Update' of an ECS is a change in the functionality and/or in the realization (implementation) of some functionality of an ECS. Updates can be done by exchanging components or modules of a (sub-)system and by exchanging (parts) of the software running on it. The former requires physical access to the product, which can be done during regular or additional maintenance services, the later can also be done ‘over the air’ only. There are three reasons for updates: The first is error correction, i.e., when an hitherto unknown functional error or a security hazard is detected, it can be corrected via an update. The second is performance or usability optimization and variation. If a function in an ECS can be implemented in different ways – for example engine control in a car can be done ‘smooth’ or ‘sportive’ or in many other ways – then a change in the implementation can be realized by updates. Updates initiated by one of these two reasons already have a highly positive impact on sustainability. However, the main benefit for sustainability stems from the third reason for updates, which is adding functionality to an ECS already in use, thus avoiding replacement of the product in favour of ‘newer, better versions’. The main challenges for updates are first, ensuring quality of the updated system. Taking safety as the major quality of an ECS again, it is difficult to analyse and guarantee all the existing safety properties of a system for the new, updated system, seeing that the updated system is already in operation and thus not available for physical tests anymore. The second challenge lies in the fact that there are typically many – sometimes even thousands – of variants of a product in operation, and the update has to fit all of them and quality has to be assured for all of them. The third challenges mostly concerns over-the-air software updates, for which safe and secure deployment has to be guaranteed.

Energy and resource efficient test procedures and equipment

Testing and validation of ECS used to be a very energy consuming task, with hundreds and thousands of physical tests taking place not only with the final product (or a prototype thereof), but also with all subsystems, components and modules of the system. The setup, maintenance and operation of the various test environments were also highly resource intensive. With the shift to virtual engineering including the use of digital twins and virtual testing (c.f. Major Challenge 1), many physical tests can be avoided. However, first, the number of physical tests needed is still fairly large, and reducing their energy and resource consumption is still an important topic. In addition, virtual engineering itself requires energy for all the computers, servers, and simulation suites that are needed for it. Any technique that lowers power consumption for computers, servers, and clouds is therefore also beneficial here, as are methods that reduce the number of test cases needed for showing quality attributes.

Ultra-low power design methods

The potential application area for ultra-low power electronic systems is very high due to the rapidly advancing miniaturisation of electronics and semiconductors, as well as the ever-increasing connectivity enabled by it. This ranges from biological implants, home automation, the condition-monitoring of materials to location-tracking of food, goods or technical devices and machines. Digital products such as radio frequency/radio frequency identification (RF/RFID) chips, nanowires, high-frequency (HF) architectures, SW architectures or ultra-low power computers with extremely low power consumption support these trends very well (see also appendix A on RISC-V). Such systems must be functional for extended periods of time with a limited amount of energy.

The ultra-low-power design methods comprise the areas of efficiency modelling and low-power optimisation with given performance profiles, as well as the design of energy-optimised computer architectures, energy-optimised software structures or special low-temperature electronics (c.f. chapters 1.1, 1.2 and 1.3). Helpful here are system-level automatic DSE (design space exploration) approaches able to fully consider energy/power issues (e.g. dark silicon, energy/power/performance trade-offs) and techniques. The design must consider the application-specific requirements, such as the functional requirements, power demand, necessary safety level, existing communication channels, desired fault tolerance, targeted quality level and the given energy demand and energy supply profiles, energy harvesting gains and, last but not least, the system's lifetime.

Exact modelling of the system behaviour of ultra-low power systems and components enables simulations to compare and analyse energy consumption with the application-specific requirements so that a global optimisation of the overall system is possible. Energy harvesting and the occurrence of parasitic effects, must also be taken into account.

2.3.2.3 Major Challenge 3: Managing complexity

2.3.2.3.1 State of the art

The new system capabilities (intelligence, autonomy, evolvability), as well as the required system properties (safety, security, reliability, trustworthiness), each considerably increase complexity (c.f. Part 3 and Sub-section 2.3.1 above). Increasingly complex environments in which these systems are expected to operate, and the increasingly complex tasks (functionalities) that these systems need to perform in this environment, are further sources of soaring system complexity. Rising complexity leads to a dramatic upsurge in the effort of designing and testing, especially for safety-critical applications where certification is usually required. Therefore, an increased time to market and increased costs are expected, and competitiveness in engineering ECS is endangered. New and improved methods and tools are needed to handle this new complexity, to enable the development and design of complex systems to fulfil all functional and non-functional requirements, and to obtain cost-effective solutions from high productivity. Three complexity-related action areas will help to master this change:

- methods to enable efficient Safety Assurance Cases
- methods and tools to increase design efficiency.
- complexity reduction methods and tools for V&V and testing.
- methods and tools for advanced architectures.

2.3.2.3.2 Vision and expected outcome

The connection of electronics systems and the fact that these systems change in functionality over their lifetime continuously drives complexity. In the design phase of new connected highly autonomous and evolvable ECS, this complexity must be handled and analysed automatically to support engineers in generating best-in-class designs with respect to design productivity, efficiency and cost reduction. New methods and tools are needed to handle this new complexity during the design, manufacturing and operations phases. These methods and tools, handling also safety related non functional requirements, should work either automatically or be recommender-based for engineers to have the complexity under control (see also the corresponding challenges in chapter 1.3. Embedded Software and Beyond).

Complexity increases the effort required, especially in the field of V&V of connected autonomous electronics systems, which depend on each other and alter over their lifetime (cf. chapter 3.1). The innumerable combinations and variety of ECS must be handled and validated. To that end, new tools and methods are required to help test engineers in creating test cases automatically, analysing testability and test coverage on the fly while optimising the complete test flow regarding test efficiency and cost. This should be achieved by identifying the smallest possible set of test cases sufficient to test all possible system behaviours. It is important to increase design efficiency and implement methods that speed up the design process of ECS. Methods and tools for X-in-the-loop simulation and testing must be developed, where X represents hardware, software, models, systems, or a combination of these. A key result of this major challenge will be the inclusion of complexity-reduction methods for future ECS-based systems into the design flows derived in Major Challenge 1, including seamless tool support, as well as modular architectures that support advanced computation methods (AI, advanced control), system improvements (updates), replacement and recycling by 2026. Building on these, modular and evolvable/extendable reference architectures and (hierarchical, open source-based) chips (i.e., RISC-V, see appendix A), modules, components, systems and platforms that support continuous system improvement, self-awareness, health and environment monitoring, and safe and secure deployment of updates, will be realised by 2029.

2.3.2.3.3 Key focus areas

Assurance Cases

Safety Assurance Cases are a commonly and successfully used method to show that a certain product possesses certain qualities, e.g., safety. Corresponding arguments for other qualities (like security, dependability, etc.) are being used as well.

Safety Assurance cases consist of a set of Safety Cases or Safety Arguments. ISO 26262 states that “the purpose of a safety case is to provide a clear, comprehensive and defensible argument, supported by evidence, that an item is free from unreasonable risk when operated in an intended context’. According to the CAE principle, Safety Arguments comprise Claims, i.e. that the system fulfills a certain safety related requirement, and Arguments, i.e., an explanation – or proof -- why a claim (or sub claim) is valid. Arguments, in turn, comprise sub-claims – for which in turn an Argument is needed -- or evidence, i.e., an analysis or test results or other findings from the implemented system that support the (sub-claim).

Safety Assurance cases are still a valid means for showing Safety (or other qualities) of an ECS-based product. However, the following non-trivial extensions to the method are needed in order to be able to handle modern and future ECS-based products:

- ODD, behavior competencies: To construct safety cases, both the operational context as well as the functionality of the system needs to be specified. While there has been considerable advancement in describing operational context by ODDs (Operational Design Domains), including first standardization activities, there still is a lot of room for further improvement. Behavior Competencies, which are used to formally describe the systems functionality, are still further behind. Together these two concepts need to be extended by means to describe reduced functional behavior and minimal risk maneuvers.
- Handling of unknowns: Both the definition of the operational context as well as the functionality of the system will contain uncertainties and unknowns (c.f. Handling of Uncertainties in Major Challenge 1). To include these in the safety argument, statistical reasoning and/or three-valued logics need to be used.
- Quality metrics and Guarantees: For highly automated (up to autonomous) systems, which may have part of their functionality implemented by Artificial Intelligence, it is sometimes infeasible to collect the evidence needed to support a certain (sub-)claim. For example, in the automotive sector for autonomous driving, test driving the number of miles required is infeasible. Sometimes, on the other hand, we do not even know what quality to measure or what guaranteed performance would be evidence to support a certain (sub-)claim. Again as an example, for AI-based object detection, we cannot interfere how much training and how many tests are required to guarantee this property. The challenge here is to find quality Metrics and Guarantees that (a) are sufficiently strong to support the safety argument and (b) can be measured or analyzed with the needed efficiency.

Methods and tools to increase design and V&V efficiency

Design efficiency is a key factor for keeping and strengthening engineering competitiveness. Design and engineering in the virtual world using simulation techniques require increasingly efficient modelling methods of complex systems and components. Virtual design methodology will be boosted by the research topics:

- XIL-testing (X-in-the-loop, with X=model, system, software, hardware,.. incl. mixed-modes
- XIL simulation techniques and tools, speed up of simulation, accuracy of simulation, multi-domain co-simulation
- Efficient modelling, test and analysis for reliable, complex systems on different abstraction levels
- Evaluation of architecture and design of the ECS SW/HW with real tests

Complexity reduction methods and tools for V&V and testing

A second way to manage complexity is the complexity-related reduction of effort during the engineering process. Complexity generates most effort in test, and V&V, ensuring compatibility and proper behaviour in networking ECS. Consistent hierarchical design and architectures, and tool-based methods to design those architectures automatically, are needed. Advanced test methods with intelligent algorithms for test termination, as well as automated metrics for testability and diagnosis (including diagnosis during run-time), must be developed and installed. This includes the following research topics:

- Recommender-based guidance in V&V process for complex ECS systems
- Automated generation of testcases from models/digital twins of ECS systems
- Test coverage calculation by means of models and testcases (coverage-driven V&V)
- Minimizing effort for V&V based on models, AI techniques
- Advanced test methods, intelligent concepts for test termination, automated metrics/tools for testability, diagnosis, and extraction of diagnostic information
- Methods and tools for consistent, hierarchical design, V&V and test
- Energy and resource efficient test procedures and equipment

Methods and tools for advanced architectures

Complexity, and also future complexity, is mainly influenced by the Architecture. Future architectures must support complex, highly connected ECS that use advanced computational methods and AI, as well as machine learning, which lead to a change of ECS over lifetime. Especially for AI (cf. chapter 2.1), this includes support for V&V of the AI method, for shielding mechanisms and other forms of fault/uncertainty detention resp. for prevention of fault propagations, and for advanced monitoring concepts, that allow deep introspection of components and modules as well as hierarchical 'flagging', merging and handling of monitoring results and detected anomalies. For this, reference architectures and platform architectures are required on all levels of the design hierarchy (for the system and SOS levels, see also the challenges "Open SoS architecture and infrastructure", "SoS interoperability" and related challenges in chapter 1.4 on System of Systems).

An additional focus of Architecture exploration and optimisation must be architectures that ease the necessary efforts for analysis, test, V&V and certification of applications. Hierarchical, modular architectures that support a divide-and-conquer approach for the design and integration of constituent modules with respect to subsystems have the potential to reduce the demand for analysis and V&V ("correct by design" approach). As integration platforms, they have to ensure interoperability of constituent ECS. For the Architecture exploration and optimisation itself, AI-based methods are needed to achieve a global optimum. Overall, holistic design approaches and tools for architectures of multi-level/multi-domain systems are the goal.

Apart from the benefits that reference architectures and platforms have at a technological level, they are also important economically. As integration platforms for solutions of different vendors, they serve as a focal point for value chain-based ecosystems. Once these ecosystems reach a certain size and market impact, the platforms can serve as the basis for corresponding "platform economies" (cf. Major Challenge "Open SoS architecture and infrastructure" in chapter 1.4). Thus, the following research topics turn out:

- Architecture exploration and optimization, including multi-aspect optimization (e.g. safety, security, comfort, functionality,…) and AI based optimization methods
- Architectures supporting advanced computation methods (AI, advanced control,…)
- Architectures and tools for non von-Neumann and neuromorphic computing
- Architectures supporting self-awareness, health and environment monitoring on all levels of the design hierarchy

- Platform and middleware architectures, also for extremely distributed, multi-layered SoS and IoT applications
- Reference architectures for continuous system improvement, i.e. across evolving system generations
- Architectures for V&V and certification, including automatic evaluation of computation and deployment decisions (i.e. on chip, edge, fog, cloud).
- Modular and evolvable/extendable architectures (supporting traceability of evolution, also supporting modular updates, replacement and recycling for a circular economy
- (SW-HW) architecture mapping (incl. resource mapping and tracing (communication, scheduling, …), incl. requirement matching and tracing

All of the above topics need to be examined and developed in conjunction with corresponding changes in design and validation tools ('Design for Target Architectures', 'Design for platforms') in order to leverage on the complexity reduction that they bring.

2.3.2.4 Major Challenge 4: Managing diversity

2.3.2.4.1 State of the art

In the ECS context, diversity is everywhere – between polarities such as analogue and digital, continuous and discrete, and virtual and physical. With the growing diversity of today's heterogeneous systems, the integration of analogue-mixed signals, sensors, micro-electromechanical systems (MEMS), chiplets, actuators and power devices, transducers and storage devices is essential. Additionally, domains of physics such as mechanical, photonic and fluidic aspects have to be considered at the system level, and for embedded and distributed software. The resulting design diversity is enormous. It requires multi-objective optimisation of systems (and SoS), components and products based on heterogeneous modelling and simulation tools, which in turn drives the growing need for heterogeneous model management and analytics. Last, but not least, a multi-layered connection between the digital and physical world is needed (for real-time as well as scenario investigations). Thus, the ability to handle this diversity on any level of the design hierarchy, and anywhere it occurs, is paramount, and a wide range of applications has to be supported.

2.3.2.4.2 Vision and expected outcome

The management of diversity has been one of Europe's strengths for many years. This is not only due to European expertise in driving More-than-Moore issues, but also because of the diversity of Europe's industrial base. Managing diversity is therefore a key competence. Research, development and innovation (R&D&I) activities in this area aim at the development of design technologies to enable the development of complex, smart and, especially, diverse systems and services. All these have to incorporate the growing heterogeneity of devices and functions, including its V&V across mixed disciplines (electrical, mechanical, thermal, magnetic, chemical and/ or optical, etc). New methods and tools are needed to handle this growing diversity during the phases of design, manufacturing and operation in an automated way. As in complexity, it is important to increase design efficiency on diversity issues in the design process of ECS. A major consequence of this challenge will be the inclusion of methods to cope with all diversity issues in future ECS-based systems, which have been introduced into the design flows derived in Major Challenge 1, including seamless tool support for engineers.

2.3.2.4.3 Key focus areas

The main R&D&I activities for this fourth major challenge are grouped into the following key focus areas.

Multi-objective design and optimisation of components and systems

The area of multi-objective optimisation of components, systems and software running on SoS comprises integrated development processes for application-wide product engineering along the value chain (cf. Part 1 and appendix A on RISC-V). It also concerns modelling, constraint management, multi-criteria, cross-domain optimisation and standardised interfaces. This includes the following research topics

- Consistent & complete Co-Design & integrated simulation of IC, package and board in the application context
- Methods and Tools to support multi-domain designs (i.e., electronic/electric and hydraulic, …) and multi paradigm designs (different vendors, modelling languages, …) and HW/SW co-design
- Advanced Design Space Exploration and iterative Design techniques, inl. Multi-aspect optimization (Performance vs. cost vs. space vs. power vs. reliability)
- Modular design of 2.5 and 3D integrated systems and chiplets and flexible substrates

Modelling, analysis, design and test methods for heterogeneous systems considering properties, physical effects and constraints

The area of modelling, analysis, design, integration and testing for heterogeneous systems considering properties, physical effects and constraints comprises the following methods and tools which all need to consider chiplet technology aspects: .

- Methods and tools for design, modelling and integration of heterogeneous systems (incl. chiplet technology)
- Hierarchical methods for hardware/software co-simulation and co-development of heterogeneous systems (multi-scale, multi-rate modelling and simulation)
- Modelling methods to take account of operating conditions, statistical scattering and system changes
- Hierarchical modelling and early assessment of critical physical effects and properties from SoC up to system level
- Analysis techniques for new circuit concepts and special operating conditions (voltage domain check, especially for start-up, floating node analysis ...)

Automation of analogue and integration of analogue and digital design methods

The area of automation of analogue and integration of analogue and digital design methods comprises the following research topics:

- Metrics for analogue/mixed signal (AMS) testability and diagnostic efficiency (including V&V & test)
- Harmonisation of methods and tooling environments for analogue, RF and digital design
- Automation of analogue and RF design – i.e. high-level description, synthesis acceleration and physical design, modularisation and the use of standardised components

Connecting the virtual and physical world of mixed domains in real environments

The area of connecting the virtual and physical worlds of mixed domains in real environments is about the following research topics:

- Advanced analysis considering the bi-directional connectivity of the virtual and physical world of ECS and its environment (including environmental modelling, multimodal simulation, simulation of (digital) functional and physical effects, emulation and coupling with real, potentially heterogenous, hardware, and integration of all of these into a continuous design and validation flow for heterogeneous systems, cf. Major Challenge 1 and 2 above).
- Novel more-than-Moore design methods and tools,
- Models and model libraries for chemical and biological systems

2.3.3 Timeline

Major Challenge	Topic / Key focus area	Short Term	Medium Term	Long Term
		2025-2029	2030-2034	2035 and beyond
MC 1: Enabling cost- and effort-efficient Design and Validation Frameworks for High Quality ECS	**Topic 1.1:** Lifecycle-aware holistic design flows	Supply-chain aware design frameworks covering the complete lifecycle of ECS, including data feedback from production, operations, and maintenance supporting multi-level, multi-paradigm continuous modelling, integration, and analysis; interoperable tool chains	Supply-chain aware design frameworks allowing 'forward optimization' (Design for optimized manufacturing / operations / recycling) as well as 'backwards optimization' (continuous design, integration and test)	
	Topic 1.2: Virtual engineering of ECS	Advanced Model-Based Design and Specification Tools for the System under Development as well as for the Environment; semi-automatic generation of design artefacts; aspect-specific Digital Twins; advanced simulation and validation tools	Domain specific Digital Twins. Accurate simulation environments enhanced by automatic validation tools; automatic generation of design artefacts; accurate simulation of environment and ECS behaviour	Full Digital Twins

	Topic 1.3 Advancing System and component design (methods and tools)	Executable models of sensors; model based design technologies, switch to probabilistic analysis of behaviour to handle uncertainties; AI supported modelling, design space exploration, and data analysis	Uncertainty – aware, residual risk based design and analysis of perception chain and prediction engines; Explainable AI-based modelling, design space exploration and data analysis	Full Explainable AI support for all phases of the development, integration and test process.
	Topic 1.4 Integration of new V&V methods	Safety Cases based on ODD and system capabilities, ODD monitoring within fail-aware ECS, scenario and test case coverage of ODD; V&V for modular Updates	Full V&V for Updates based on Digital Twins. V&V for systems with shielded or guarded AI based components	V&V for systems with AI-based components; certification at run-time (for known environments and restricted updates)
MC 2: Enabling Sustainable Design for Sustainability	**Topic 2.1:** Lifecycle aware Design Optimizations	Requirement engineering for 'Design for X', specification, implementation and testing techniques; health monitoring of system components. AI supported analysis of health data	Resolution techniques for conflicting requirements; AI based analysis of health data	
	Topic 2.2: Updates	Development and secure deployment of safe updates based on selected data from the field; Variant Management for Updates, V&V for modular updates	Full V&V for variant aware Updates based on Domain-specific Digital Twins within virtual engineering frameworks	

	Topic 2.3: Energy and resource efficient test procedures and equipment	c.f. Topic 1.2 'Virtual Engineering for ECS' above as well as Major Challenge 1 of Chapter 2.1		
	Topic 2.4: Ultra-low power design methods	System level DSE tools supporting energy aware design decisions; energy aware modelling of system behaviour, energy aware simulation of Ultra-low-power systems	Increased accuracy of Modelling and simulation techniques considering fault tolerances, targeted quality level, etc.	
MC 3: Managing complexity	**Topic 3.1:** Assurance Cases	Safety Assurance Cases based on ODD and system capabilities/behaviour competencies, ODD monitoring within fail-aware ECS, scenario and test case coverage of ODD; Quality metrics, Quality guaranties	Handling of unknown scenarios and system behaviour uncertainties beyond Minimum Risk Manoeuvres; Efficient validation of quality metrics and guarantees based on statistical analysis.	Assurance Cases for fully autonomous systems
	Topic 3.2: Methods and tools to increase design and V&V efficiency	Fully hierarchical XIL testing with clearly defined relations between abstraction levels, multi-domain, multi-objective simulation; validation of simulation results by physical testing; AI Support (c.f. Topic 1.3)	Accurate and fast simulation of system behaviour and environment; minimization of the need for physical tests; AI Support (c.f. Topic 1.3)	Fully simulation based V&V using Digital Twins; AI Support (c.f. Topic 1.3)

	Topic 3.3: Complexity reduction methods and tools for V&V and testing	Automated generation of test cases from models/digital twins of ECS systems; coverage-driven V&V, AI supported V&V and testing	Recommender-based guidance in V&V process for complex ECS systems; Explainable AI-based V&V	Complexity reduction and V&V techniques fully based on explainable AI
	Topic 3.4: Methods and tools for advanced architectures	Hierarchical, modular architectures for V&V and certification supporting AI and advanced control, multi-aspect design space exploration and multi-objective optimizations for Architecture Design, Reference Architectures supporting continuous system improvement	Modular and evolvable/extendable Reference Architectures and Platforms, design tools for architectures for multi-level, multi-domain systems	Architectures and Tools for new technologies, holistic design approaches and tools for architectures for multi-level, multi-domain systems
MC 4: Managing diversity	**Topic 4.1:** Multi-objective design and optimisation of components and systems	Methods and Tools to support multi-domain and multi paradigm designs, as well as HW/SW co-design; Modular design of 2.5 and 3D integrated systems and chiplets and flexible substrates	Consistent & complete Co-Design & integrated simulation of IC, package and board in the application context; Advanced Design Space Exploration and iterative Design techniques, incl. Multi-aspect optimization	

	Topic 4.2: Modelling, analysis, design and test methods for heterogeneous systems considering properties, physical effects and constraints	Methods and tools for design, modelling and integration of heterogeneous systems (incl. chiplet technology); Hierarchical methods for HW/SW co-simulation and co-development of heterogeneous systems	Modelling methods to take account of operating conditions, statistical scattering and system changes; Hierarchical modelling and early assessment of critical physical effects and properties from SoC up to system level; Analysis techniques for new circuit concepts and special operating conditions	
	Topic 4.3: Automation of analogue and integration of analogue and digital design methods	Harmonisation of methods and tooling environments for analogue, RF and digital design; Automation of analogue and RF design – i.e. high-level description, synthesis acceleration and physical design, modularisation and the use of standardised components	Metrics for analogue/mixed signal (AMS) testability and diagnostic efficiency (including V&V & test)	
	Topic 4.4: Connecting the virtual and physical world of mixed domains in real environments	Novel more-than-Moore design methods and tools	Advanced analysis considering the bi-directional connectivity of the virtual and physical world of ECS and its environment (cf. Major Challenge 1 and 2 above); Models and model libraries for chemical and biological systems	

2.4

Cross-Sectional Technologies

QUALITY, RELIABILITY, SAFETY AND CYBERSECURITY

2.4. QUALITY, RELIABILITY, SAFETY AND CYBERSECURITY

2.4.1. SCOPE

Modern technologies and new digitised services are key to ensuring the stable growth and development of the European Union and its society. These new technologies are largely based on smart electronic components and systems (ECS). Highly automated or autonomous transportation systems, improved healthcare, industrial production, information and communication networks, and energy grids all depend on the availability of electronic systems. The main societal functions[1] and critical infrastructure are governed by the efficient accessibility of smart systems and the uninterrupted availability of services.

Ensuring the reliability, safety and security of ECS is a Major Challenge since the simultaneous demand for increased functionality and continuous miniaturisation of electronic components and systems causes interactions on multiple levels. A degraded behaviour in any of these dimensions (quality, reliability, safety, and security) or an incorrect integration among them, would affect vital properties and could cause serious damage. In addition, such shortcomings in safety, reliability and security might even outweigh the societal and individual benefits perceived by users, thus lowering trust in, and acceptance of, the technologies.
These topics and features constitute the core of this Chapter, which addresses these complex interdependencies by considering input from, and necessary interaction between, major disciplines. Moreover, quality, reliability, safety and cybersecurity of electronic components and systems are, and will be, fundamental to digitised society (see Figure 2.4.1). In addition, the tremendous increase of computational power and reduced communication latency of components and systems, coupled with hybrid and distributed architectures, impose to rethink many "traditional" approaches and expected performances towards safety and security, exploiting AI and ML (machine learning).

In practice, ensuring reliability, safety, and security of ECS is part of the Design, Implementation, and Validation/Testing process of the respective manufacturers and – for reasons of complexity and diversity/heterogeneity of the systems – must be supported by (analysing and testing) tools. Thus, the techniques described in Chapter 2.3 (Architecture and Design: Method and Tools) are complementary to the techniques presented here: in that Chapter, corresponding challenges are described from the design process viewpoint, whereas here we focus on a detailed description of the challenges concerning reliability, safety, and security within the levels of the design hierarchy.

---

1 Vital societal functions: services and functions for maintaining the functioning of a society. Societal functions in general: various services and functions, public and private, for the benefit of a population and the functioning of society.

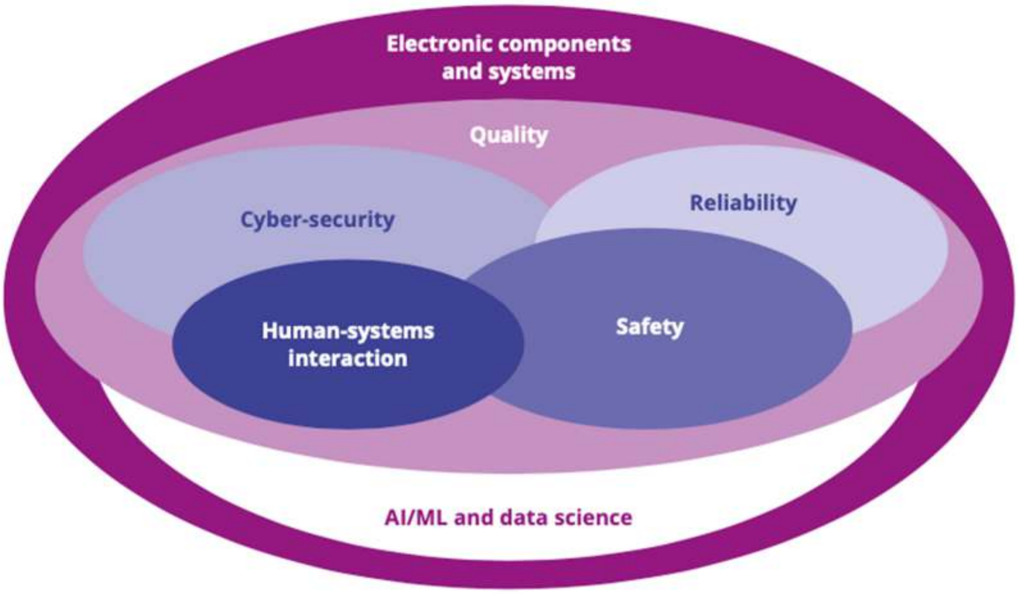


Figure 2.4.1 - Role of quality, reliability, safety and cybersecurity of electronic components and systems for digitalisation.

2.4.2. Major Challenges

To introduce the topic presented in this Chapter, we first present some definitions that will be useful to clarify the concepts described in the **Major Challenges**.

- Production quality: often defined as "the ability of a system being suitable for its intended purpose while satisfying customer expectations", this is a very broad definition that basically includes everything. Another widely used definition is "the degree a product meets requirements in specifications" – but without defining the underlying specifications, the interpretation can vary a lot between different stakeholders. Therefore, in this Chapter quality will be defined "as the degree to which a product meets requirements in specifications that regulate how the product should be designed and manufactured, including environmental stress screening (such as burn-in) but no other type of testing". In this way, reliability, dependability and cybersecurity, which for some would be expected to be included under quality, will be treated separately.
- Reliability: this is the ability or the probability, respectively, of a system or component to function as specified under stated conditions for a specified time.
- Prognostics and health management: a method that permits the assessment of the reliability of the product (or system) under its application conditions.
- Functional safety: the ability of a system or piece of equipment to control recognized hazards to achieve an acceptable level of risk, such as to maintain the required minimum level of operation even in the case of likely operator errors, hardware failures and environmental changes to prevent physical injuries or damages to the health of people, either directly or indirectly.
- Dependability: according to IEC 60050-192:2015, dependability (192-01-22) is the ability of an item to perform as and when required. An item here (192-01-01) can be an individual part, component, device, functional unit, equipment, subsystem or

system. Dependability includes availability (192-01-23), reliability (192-01-24), recoverability (192-01-25), maintainability (192-01- 27) and maintenance support performance (192-01-29), and in some cases other characteristics such as durability (192-01-21), safety and security. A more extensive description of dependability is available from the IEC technical committee on dependability (IEC TC 56).

- Safety: freedom from unacceptable risk of harm [CENELEC 50126].
- Security: measures can provide controls relating to physical security (control of physical access to computing assets) or logical security (capability to login to a given system and application) (IEC 62443-1-1):
  - measures taken to protect a system;
  - condition of a system that results from the establishment and maintenance of measures to protect the system;
  - condition of system resources being free from unauthorized access, and from unauthorized or accidental change, destruction or loss;
  - capability of a computer-based system to provide adequate confidence that unauthorized persons and systems can neither modify the software and its data nor gain access to the system functions, and yet ensure that this is not denied to authorized persons and systems;
  - prevention of illegal or unwanted penetration of, or interference with, the proper and intended operation of an industrial automation and control system.
- Cybersecurity: the protection of information against unauthorized disclosure, transfer, modification or destruction, whether accidental or intentional (IEC 62351-2).
- Robust root of trust systems: these are based on cryptographic functionalities that ensure the authenticity and integrity of the hardware and software components of the system, with assurance that it is resilient to logical and physical attacks.
- Emulation and Forecasting: cybersecurity evolution in parallel to increasing computation power and hybrid threats mixing geopolitical, climate change and any other external threats impose to anticipate the horizon of resilience, safety and security of systems forecasting attacks and incidents fast evolution.

Five Major Challenges have been identified:

- **Major Challenge 1**: ensuring HW quality and reliability.
- **Major Challenge 2**: ensuring dependability in connected software.
- **Major Challenge 3**: ensuring cyber-security and privacy.
- **Major Challenge 4**: ensuring of safety and resilience.
- **Major Challenge 5**: human systems integration.

2.4.4.1 Major challenge 1: Ensuring HW quality and reliability

2.4.4.1.1. State of the art

With the ever-increasing complexity and demand for higher functionality of electronics, while at the same time meeting the demands of cutting costs, lower levels of power consumption and miniaturization in integration, hardware development cannot be decoupled from software development. Specifically, when assuring reliability, separate hardware development and testing according to the second-generation reliability methodology (design for reliability, DfR) is not sufficient to ensure the reliable function of the ECS. A third-generation reliability methodology must be introduced to meet these challenges. For the electronic smart systems used in future highly automated and autonomous systems, a next generation of reliability is therefore required. This new generation of reliability assessment will introduce in situ monitoring of the state of health on both a local (e.g. IC packaging) and system level. Hybrid prognostic and health management (PHM) supported by Artificial Intelligence (AI) is the key methodology here. This marks the main difference between the second and the third generation. DfR concerns the total lifetime of a full population of systems under anticipated service conditions and its statistical characterization. PHM, on the other hand, considers the degradation of the individual system in its actual service conditions and the estimation of its specific remaining useful life (RUL).

2.4.4.1.2. Vision and expected outcome

Since embedded systems control so many processes, the increased complexity by itself is a reliability challenge. Growing complexity makes it more difficult to foresee all dependencies during design. It is impossible to test all variations, and user interfaces need greater scrutiny since they have to handle such complexity without confusing the user or generating uncertainties.

The trend towards interconnected, highly automated and autonomous systems will change the way we own products. Instead of buying commodity products, we will instead purchase personalized services. The vision of **Major challenge 1** is to provide the requisite tools and methods for novel ECS solutions to meet ever- increasing product requirements and provide availability of ECS during use in the field. Therefore, availability will be the major feature of ECS. Both the continuous improvement of existing methods (e.g., DfR) and development of the new techniques (PHM) will be the cornerstone of future developments in ECS (see also Challenges 1 and 2, and especially the key focus areas on lifecycle-aware holistic design flows in Chapter 2.3 **Architecture and Design: Methods and Tools**). The main focus of Major challenge 1 will circulate around the following topics.

- Digitization, by improving collaboration within the supply chain to introduce complex ECS earlier in the market.
- Continuous improvement of the DfR methodology through simultaneous miniaturization and increasing complexity.
- Model-based design is a main driver of decreasing time-to-market and reducing the cost of products.
- Availability of the ECS for highly automated and autonomous systems will be successfully introduced in the market based on PHM.
- Data science and AI will drive technology development and pave the way for PHM implementation for ECS.
- AI and PHM based risk management.

2.4.4.1.3. Key focus areas

2.4.4.1.3.1 Quality: In situ and real-time assessments

Inline inspection and highly accelerated testing methods for quality and robustness monitoring during production of ECS with ever-increasing complexity and heterogeneity for demanding applications should increase the yield and reduce the rate of early fails (failures immediately following the start of the use period).

- Controlling, beyond traditional approaches, the process parameters in the era of Industry 4.0 to minimize deviations and improve quality of key performance indicators (KPIs).
- Process and materials variabilities will have to be characterized to quantify their effects on hardware reliability, using a combination of empirical studies, fundamental RP models and AI approaches.
- Advanced/smart monitoring of process output (e.g., measuring the 3D profile of assembled goods) for the detection of abnormities (using AI for the early detection of standard outputs).
- Early detection of potential yield/reliability issues by simulation-assisted design for assembly/design for manufacturing (DfM/DfA) as a part of virtual prototyping.

2.4.4.1.3.2 Digitization: A paradigm shift in the fabrication of ECS from supplier/customer to partnership

Digitization is not possible without processing and exchange data between partners.

- Involving European stakeholders to resolve the issue of data ownership:
    - Create best practices and scalable workflows for sharing data across the supply chain while maintaining intellectual property (IP).
    - Standardize the data exchange format, procedures and ownership, and create an international legal framework.
    - Conceive and validate business models creating economic incentives and facilitating sharing data, and machine learning algorithms dealing with data.
- Handling and interpreting big data:
    - Realise consistent data collection and ground truth generation via annotation/labelling of relevant events.
    - Create and validate a usable and time-efficient workflow for supervised learning.
    - Standardized model training and model testing process.
    - Standardized procedures for model maintenance and upgrade.
- Make a link between data from Industry 4.0 and model-based engineering:
    - Derive working hypotheses about system health.
    - Validate hypothesis and refine physics-based models.
    - Construct data models-based embedding (new) domain knowledge derived from model-based engineering.

- Identify significant parameters that must be saved during production to be re-used later for field-related events, and vice versa – i.e., feed important insights derived from field data (product usage monitoring) into design and production. This is also mandatory to comply with data protection laws.
- Evaluate methods for the indirect characterization of ECS using end-of-line test data.
- Wafer fabrication (pre-assembly) inline and offline tests for electronics, sensors and actuators, and complex hardware (e.g. multicore, graphics processing unit, GPU) that also cover interaction effects such as heterogeneous 3D integration and packaging approaches for advanced technologies nodes (e.g. thin dice for power application – dicing and grinding).

2.4.4.1.3.3 Reliability: Tests and modelling

Continuous improvement of physics of failure (PoF) based methodologies combined with new data-driven approaches: tests, analyses and degradation, and lifetime models (including their possible reconfiguration):

- Identifying and adapting methodology to the main technology drivers.
- Methods and equipment for dedicated third-level reliability assessments (first level: component; second level: board; third level: system with its housing, e.g. massive metal box), as well as accounting for the interactions between the hierarchy levels (element, device, component, sub-module, module, system, application).
- Comprehensive understanding of failure mechanisms, lifetime prediction models (including multi-loading conditions), continuously updating for new failure mechanisms related to innovative technologies (advanced complementary metal–oxide–semiconductor (CMOS), µ-fluidics, optical input/output (I/O), 3D printing, wide bandgap technologies, etc). New materials and production processes (e.g. 3D printing, wide bandgap technologies, etc), and new interdisciplinary system approaches and system integration technologies (e.g. µ-fluidics, optical input/output (I/O), etc).
- Accelerated testing methods (e.g. high temperature, high power applications) based on mission profiles and failure data (from field use and tests):
    - Use field data to derive hypotheses that enable improved prioritization and design of testing.
    - Usage of field, PHM and test data to build models for ECS working at the limit of the technology as accelerating testing is limited.
- Standardize the format of mission profiles and the procedure on how mission profiles are deducted from multimodal loading.
- Design to field – better understanding of field conditions through standardized methodology over supply chain using field load simulator.
- Understanding and handling of new, unforeseen and unintended use conditions for automated and autonomous systems.
- Embedded reliability monitoring (pre-warning of deterioration) with intelligent feedback towards autonomous system(s).
- Identification of the 10 most relevant field-related failure modes based on integrated mission profile sensors.
- Methods to screen out weak components with machine learning (ML) based on a combination of many measured parameters or built-in sensor data.

- New standards/methodologies/paradigms that evaluate the “ultimate” strength of systems – i.e. no longer test whether a certain number of cycles are “pass”, but go for the limit to identify the actual safety margin of systems, and additionally the behavior of damaged systems, so that AI can search for these damage patterns.
- Digital twin software development for reliability analysis of assets/machines, etc.
- Comprehensive understanding of the SW influence on HW reliability and its interaction:
    - SW Reliability: start using maturity growth modelling techniques, develop models and gather model parameters.
    - SW/HW Reliability modelling: find ways as to combine the modelling techniques (in other words: scrunch the different time domains).
    - SW/HW Reliability testing: find ways as to test systems with software and find the interaction failure modes.

2.4.4.1.3.4 Design for reliability: Virtual reliability assessment prior to the fabrication of physical HW

Approaches for exchanging digital twin models along the supply chain while protecting sensitive partner IP and adaptation of novel standard reliability procedures across the supply chain.

- Digital twin as main driver of robust ECS system:
    - Identifying main technology enablers.
    - Development of infrastructure required for safe and secure information flow.
    - Development of compact PoF models at the component and system level that can be executed in situ at the system level – metamodels as the basis of digital twins.
    - Training and validation strategies for digital twins.
    - Digital twin-based asset/machine condition prediction.
- Electronic design automation (EDA) tools to bridge the different scales and domains by integrating a virtual design flow.
- Virtual design of experiment as a best practice at the early design stage.
- Realistic material and interface characterization depending on actual dimensions, fabrication process conditions, ageing effects, etc., covering all critical structures, generating strength data of interfaces with statistical distribution.
- Mathematical reliability models that also account for the interdependencies between the hierarchy levels (device, component, system).
- Mathematical modelling of competing and/or superimposed failure modes.
- New model-based reliability assessment in the era of automated systems.
- Development of fully harmonized methods and tools for model-based engineering across the supply chain:
    - Material characterization and modelling, including effects of ageing.
    - Multi-domain physics of failure simulations.
    - Reduced modelling (compact models, metamodels, etc.).
    - Failure criteria for dominant failure modes.
    - Verification and validation techniques.

- Standardization as a tool for model-based development of ECS across the supply chain:
  - Standardization of material characterization and modelling, including effects of ageing.
  - Standardization of simulation-driven design for excellence (DfX).
  - Standardization of model exchange format within supply chain using functional mock-up unit (FMU) and functional mock-up interface (FMI) (and also components).
  - Simulation data and process management.
  - Initiate and drive standardization process for above-mentioned points.
  - Extend common design and process failure mode and effect analysis (FMEA) with reliability risk assessment features (“reliability FMEA”).
  - Generic simulation flow for virtual testing under accelerated and operational conditions (virtual “pass/fail” approach).
- Automation of model build-up (databases of components, materials).
- Use of AI in model parametrization/identification, e.g. extracting material models from measurement.
- Virtual release of ECS through referencing.

2.4.4.1.3.5 Prognostics and health management of ECS: Increase in functional safety and system availability

- Self-monitoring, self-assessment and resilience concepts for automated and autonomous systems based on the merger of PoF, data science and ML for safe failure prevention through timely predictive maintenance.
- Self-diagnostic tools and robust control algorithms validated by physical fault-injection techniques (e.g. by using end-of-life (EOL) components).
- Hierarchical and scalable health management architectures and platforms, integrating diagnostic and prognostic capabilities, from components to complete systems.
- Standardized protocols and interfaces for PHM facilitating deployment and exploitation.
- Monitoring test structures and/or monitor procedures on the component and module levels for monitoring temperatures, operating modes, parameter drifts, interconnect degradation, etc.
- Identification of early warning failure indicators and the development of methods for predicting the remaining useful life of the practical system in its use conditions.
- Development of schemes and tools using ML techniques and AI for PHM.
- Implementation of resilient procedures for safety-critical applications.
- Big sensor data management (data fusion, find correlations, secure communication), legal framework between companies and countries).
- Distributed data collection, model construction, model update and maintenance.
- Concept of digital twin: provide quality and reliability metrics (key failure indicator, KFI).
- Using PHM methodology for accelerated testing methods and techniques.
- Development of AI-supported failure diagnostic and repair processes for improve field data quality.
- AI-based asset/machine/robot life extension method development based on PHM.

- AI-based autonomous testing tool for verification and validation (V&V) of software reliability.
- Lifecycle management – modeling of the cost of the lifecycle.

2.4.4.2. Major Challenge 2: Ensuring dependability in connected software

2.4.4.2.1 State of the art

Connected software applications such as those used on the Internet of Things (IoT) differ significantly in their software architecture from traditional reliable software used in industrial applications. The design of connected IoT software is based on traditional protocols originally designed for data communications for PCs accessing the internet. This includes protocols such as transmission control protocol/internet protocol (TCP/IP), the re-use of software from the IT world, including protocol stacks, web servers and the like. This also means the employed software components are not designed with dependability in mind, as there is typically no redundancy and little arrangements for availability. If something does not work, end-users are used to restarting the device. Even if it does not happen very often, this degree of availability is not sufficient for critical functionalities, and redundancy hardware and back-up plans in ICT infrastructure and network outages still continue to occur. Therefore, it is of the utmost importance that we design future connected software that is conceived either in a dependable way or can react reliably in the case of infrastructure failures to achieve higher software quality.

2.4.4.2.2 Vision and expected outcome

The vision is that networked systems will become as dependable and predictable for end-users as traditional industrial applications interconnected via dedicated signal lines. This means that the employed connected software components, architectures and technologies will have to be enriched to deal with dependability for their operation. Future dependable connected software will also be able to detect in advance if network conditions change – e.g. due to foreseeable transmission bottlenecks or planned maintenance measures. If outages do happen, the user or end application should receive clear feedback on how long the problem will last so they can take potential measures. In addition, the consideration of redundancy in the software architecture must be considered for critical applications. The availability of a European ecosystem for reliable software components will also reduce the dependence on current ICT technologies from the US and China.

2.4.4.2.3 Key focus areas

2.4.4.2.3.1Dependable connected software architectures

In the past, reliable and dependable software was always directly deployed on specialised, reliable hardware. However, with the increased use of IoT, edge and cloud computing, critical software functions will also be used that are completely decoupled from the location of use (e.g. in use cases where the police want to stop self-driving cars from a distance):

- Software reliability in the face of infrastructure instability.

- Dependable edge and cloud computing, including dependable and reliable AI/ML methods and algorithms.
- Dependable communication methods, protocols and infrastructure.
- Formal verification of protocols and mechanisms, including those using AI/ML.
- Monitoring, detection and mitigation of security issues on communication protocols.
- Quantum key distribution ("quantum cryptography").
- Increasing software quality by AI-assisted development and testing methods.
- Infrastructure resilience and adaptability to new threats.
- Secure and reliable over-the-air (OTA) updates.
- Using AI for autonomy, network behaviour and self-adaptivity.
- Dependable integration platforms.
- Dependable cooperation of System of Systems (SoS).

This Major Challenge is tightly interlinked with the cross-sectional technology of 2.2 Connectivity Chapter, where the focus is on innovative connectivity technologies. The dependability aspect covered within this challenge is complementary to that chapter since dependability and reliability approaches can also be used for systems without connectivity.

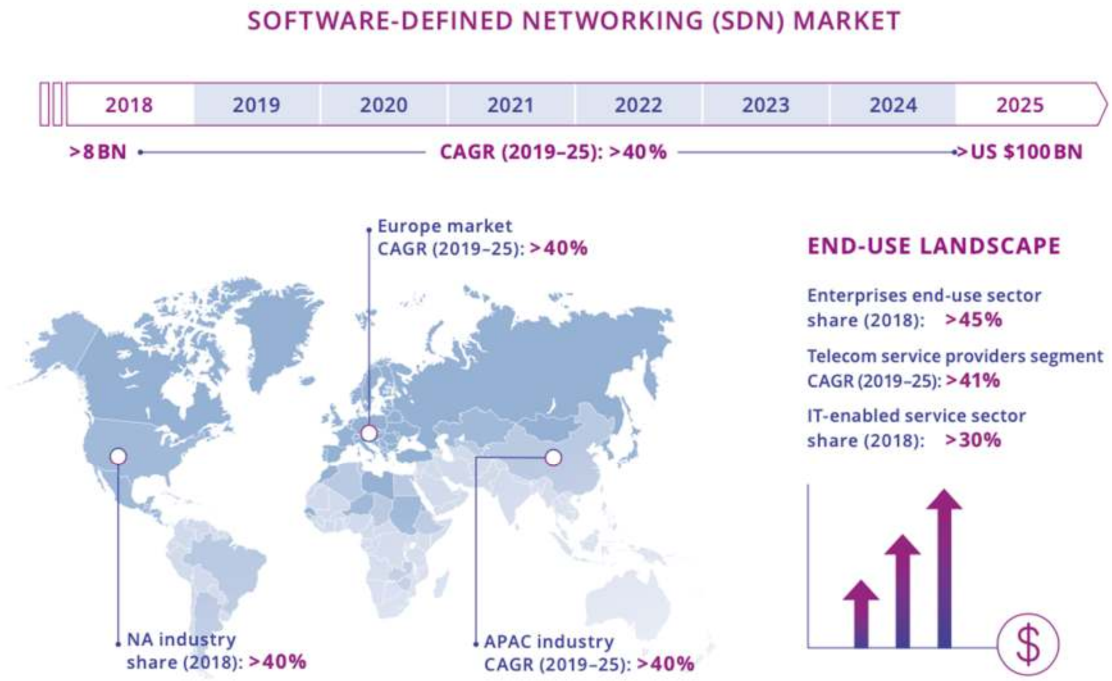


Figure 2.4.2 - Software-defined networking (SDN) market size by 2025 (Source: Global Markets Insight, Report ID GMI2395, 2018)

2.4.4.2.3.2Dependable softwarisation and virtualisation technologies

Changing or updating software by retaining existing hardware is quite common in many industrial domains. However, keeping existing reliable software and changing the underlying hardware is difficult, especially for critical applications. By decoupling software functionalities from the underlying hardware, softwarisation and virtualisation are two disruptive paradigms that can bring enormous flexibility and thus promote strong growth in the market (see Figure

2.4.2.4.2). However, the softwarisation of network functions raises reliability concerns, as they will be exposed to faults in commodity hardware and software components:

- Software-defined radio (SDR) technology for highly reliable wireless communications with higher immunity to cyber-attacks.
- Network functions virtualisation infrastructure (NFVI) reliability.
- Reliable containerisation technologies.
- Resilient multi-tenancy environments.
- AI-based autonomous testing for V&V of software reliability, including the software-in-the-loop (SiL) approach.
- Testing tools and frameworks for V&V of AI/ML-based software reliability, including the SiL approach.

2.4.4.2.3.3Combined SW/HW test strategies

Unlike hardware failures, software systems do not degrade over time unless modified. The most effective approach for achieving higher software reliability is to reduce the likelihood of latent defects in the released software. Mathematical functions that describe fault detection and removal phenomenon in software have begun to emerge. These software reliability growth models (SRGM), in combination with Bayesian statistics, need further attention within the hardware-orientated reliability community over the coming years:

- HW failure modes are considered in the software requirements definition.
- Design characteristics will not cause the software to overstress the HW, or adversely change failure-severity consequences on the occurrence of failure.
- Establish techniques that can combine SW reliability metrics with HW reliability metrics.
- Develop efficient (hierarchical) test strategies for combined SW/HW performance of connected products.

Dependability in connected software is strongly connected with other chapters in this document. In particular, additional challenges are handled in following chapters:

- Chapter 1.3 Embedded Software and Beyond: Major Challenge 1 (MC1) efficient engineering of software; MC2 continuous integration of embedded software; MC3 lifecycle management of embedded software; and MC6 Embedding reliability and trust.
- Chapter 1.4 System of Systems: MC1 SoS architecture; MC4 Systems of embedded and cyber-physical systems engineering; and MC5 Open system of embedded and cyber-physical systems platforms.
- Chapter 2.1 Edge Computing and Embedded Artificial Intelligence: MC1: Increasing the energy efficiency of computing systems.
- Chapter 2.2 Connectivity: MC4: Architectures and reference implementations of interoperable, secure, scalable, smart and evolvable IoT and SoS connectivity.
- Chapter 2.3 Architecture and Design: Method and Tools: MC3: Managing complexity.

2.4.4.3. Major Challenge 3: Ensuring cyber-security and privacy

2.4.4.3.1 State of the art

We have witnessed a massive increase in pervasive and potentially connected digital products in our personal, social and professional spheres, enhanced by new features of 5G networks and beyond. Connectivity provides better flexibility and usability of these products in different sectors, with a tremendous growth of sensitive and valuable data. Moreover, the variety of deployments and configuration options and the growing number of sub-systems changing in dynamicity and variability increase the overall complexity. In this scenario, new security and privacy issues have to be addressed, also considering the continuously evolving threat landscape. New approaches, methodologies and tools for risk and vulnerability analysis, threat modelling for security and privacy, threat information sharing and reasoning are required. Artificial intelligence (e.g., machine learning, deep learning and ontology) not only promotes pervasive intelligence supporting daily life, industrial developments, personalisation of mass products around individual preferences and requirements, efficient and smart interaction among IoT in any type of services, but It also fosters automation, to mitigate such complexity and avoid human mistakes.

Embedded and distributed AI functionality is growing at speed in both (connected) devices and services. AI-capable chips will also enable edge applications allowing decisions to be made locally at device level. Therefore, resilience to cyber-attacks is of utmost importance. AI can have a direct action on the behaviour of a device, possibly impacting its physical life inducing potential safety concerns. AI systems rely on software and hardware that can be embedded in components, but also on the set of data generated and used to make decisions. Cyber-attacks, such as data poisoning or adversarial inputs, could cause physical harm and/or also violate privacy. The development of AI should therefore go hand in hand with frameworks that assess security and safety to guarantee that AI systems developed for the EU market are safe to use, trustworthy, reliable and remain under control (C.f. Chapter 1.3 "Embedded Software and beyond" for quality of AI used in embedded software when being considered as a technology interacting with other software components).

Approaches for providing continuously evaluation of the compliance of Systems of Systems with given security standards (e.g., IEC 62443, which uses technical security controls*) will allow for the guarantee of a homogenous level of security amongst a multi-stakeholder ecosystem, challenging tech giants with platform providing overall levels of security but often resulting in vendor lock-ins. Some initial approaches resulted in products like Lynis (https://cisofy.com/lynis/) which provide continuous evaluation of some (Lynis) product specific policies. However, the rise of powerful language models and code generation may allow for a dynamic creation of evaluation machinery to support evaluation of compliance against any given standard.

The combination of composed digital products and AI highlights the importance of trustable systems that weave together privacy and cybersecurity with safety and resilience. Automated vehicles, for example, are adopting an ever-expanding combination of Advanced Driver Assistance Systems (ADAS) developed to increase the level of safety, driving comfort exploiting different type of sensors, devices and on-board computers (sensors, Global Positioning System (GPS), radar, lidar, cameras, on-board computers, etc.). To complement ADAS systems, Vehicle to X (V2X) communication technologies are gaining momentum. Cellular based V2X communication provides the ability for vehicles to communicate with other vehicle and infrastructure and environment around them, exchanging both basic safety messages to avoid collisions and, according to the 5g standard evolutions, also high throughput sensor sharing, intent trajectory sharing, coordinated driving and autonomous driving. The connected autonomous vehicle scenarios offer many advantages in terms of safety, fuel consumption and CO_2 emissions reduction, but the increased connectivity, number of devices and automation, expose those systems to several crucial cyber and privacy threats, which must be addressed and mitigated.

Autonomous vehicles represent a truly disruptive innovation for travelling and transportation, and should be able to warrant confidentiality of the driver's and vehicle's information. Those vehicles should also avoid obstacles, identify failures (if any) and mitigate them, as well as prevent cyber-attacks while staying safely operational (at reduced functionality) either through human-initiated intervention, by automatic inside action or remotely by law enforcement in the case of any failure, security breach, sudden obstacle, crash, etc.

In the evoked scenario the main cybersecurity and privacy challenges deal with:

- Interoperable security and privacy management in heterogeneous systems including cyber-physical systems, IoT, virtual technologies, clouds, communication networks, autonomous systems.
- Real time monitoring and privacy and security risk assessment to manage the dynamicity and variability of systems.
- Developing novel privacy preserving identity management and secure cryptographic solutions.
- Novel approaches to hardware security vulnerabilities and other system weaknesses as - for instance – Spectre and Meltdown or side channel attacks.
- Developing new approaches, methodologies and tools empowered by AI in all its declinations (e.g., machine learning, deep learning, ontology).
- Investigating a deep verification approach towards also open-source hardware in synergy and implementing the security by-design paradigm.
- Investigating the interworking among safety, cybersecurity, trustworthiness, privacy and legal compliance of systems.
- Evaluating the impact in term of sustainability and green deal of the adopted solutions.

2.4.4.3.2 Vision and expected outcome

The cornerstone of our vision rests on the following four pillars. First, a robust root of trust system, with unique identification enabling security without interruption from the hardware level right up to the applications, including AI, involved in the accomplishment of the system's mission in dynamic unknown environments. This aspect has a tremendous impact on mission critical systems with lots of reliability, quality and safety & security concerns. Second, protection of the EU citizen's privacy and security while at the same keeping usability levels and operation in a competitive market where also industrial Intellectual Protection should be considered. Third, the proposed technical solutions should contribute to the green deal ambition, for example by reducing their environmental impact. Finally, proof-of-concept demonstrators that are capable of simultaneously guaranteeing (a given level of) security and (a given level of) privacy, as well as potentially evolving in-reference designs that illustrate how practical solutions can be implemented (i.e. thereby providing guidelines to re-use or adapt).

End to end encryption of data, both in transit and at rest is kept to effectively protect privacy and security. The advent of quantum computing technology introduces new risks and threats, since attacks using quantum computing may affect traditional cryptographic mechanisms. New quantum safe cryptography is required, referring both to quantum cryptography and post quantum cryptography with standard crypto primitives.

Also, the roadmap for Open Source HW/SW & RISC-V IP blocks will open the path to domain-focused processors or domain-specific architectures (for instance, but not limited to the "chiplet-based approach"), which may lead to new approaches to cybersecurity and safety functions or implementation as well as new challenges and vulnerabilities that must be analysed.

Putting together seamlessly security and privacy requirements is a difficult challenge that also involves some non-technical aspects. The human factor can often cause security and privacy concerns, despite of technologically advanced tools and solutions. Another aspect relates to security certification versus certification cost. A certification security that does not mitigate the risks and threats, increases costs with minimal benefits. Therefore, all techniques and methodology to reduce such a cost are in the scope of the challenge.

In light of this scenario, this Major Challenge aims at contributing to the European strategic autonomy plan in terms of cybersecurity, digital trustworthiness and the protection of personal data.

2.4.4.3.3 Key focus areas

2.4.4.3.3.1 Trustworthiness

Digital Trust is mandatory in a global scenario, based on ever-increasing connectivity, data and advanced technologies. Trustworthiness is a high-level concern including not only privacy and security issues, but also safety and resilience and reliability. The goal is a robust, secure, and privacy preserving system that operates in a complex ecosystem without interruption, from the hardware level up to applications, including systems that may be AI-enabled. This challenge calls for a multidisciplinary approach, spanning across technologies, regulations, compliance, legal and economic issues. To this end, the main expected outcomes can be defined as:

- Defining different methods and techniques of trust for a system, and proving compliance to a security standard via certification schemes.

- Defining methods and techniques to ensure trustworthiness of AI algorithms, included explainable (XAI) (cfr. Chapter 2.1 "Edge Computing and Embedded Artificial Intelligence")
- Developing methodologies and techniques from hardware trustworthy to software layers trustworthy (cfr. Chapter 1.3 "Embedded Software and Beyond" and Chapter 1.4 "System of Systems").
- Defining methods and tools to support the composition and validation of certified parts addressing multiple standards (cfr. Chapter 1.4 "System of Systems" and Chapter 2.3 "Architecture and Design: Methods and Tools").
- Definition and future consolidation of a framework providing guidelines, good practices and standards oriented to trust.
- Enhancing current tools and procedures for safety and security verification and certification for Open-Source Hardware/Software.
- Architectures that provide mitigation, remediation and restoration against physical and software cyber threats ensuring integrity in Data, Software and Systems.

2.4.4.3.3.2 Security and privacy-by-design

The main expected outcome is a set of solutions to ensuring the protection of personal data in the embedded AI and data-driven digital economy against potential cyber-attacks:

- Ensuring cybersecurity and privacy of systems in the Edge to cloud continuum, via efficient automated verification and audits, as well as recovery mechanism (cfr. Chapter 1.4 "System of Systems" and Chapter 2.3 "Architecture and Design: Method and Tools").
- Ensuring performance in AI-driven algorithms (which needs considerable data) while guaranteeing compliance with European privacy standards (e.g., general data protection regulation - GDPR).
- Establishing a cybersecurity and privacy-by-design European data strategy to promote data sovereignty.
- Establishing Quantum-Safe Cryptography Modules.
- Establishing a transparency security approach toward Open-Source Hardware/Software Architecture.

2.4.4.3.3.3 Ensuring both safety and security properties

The main expected outcome is to ensure compatibility, adequacy and coherence in the joint use of the promoted security solutions, and the safety levels required by the system or its components:

- Maintaining the nominal or degraded system safe level behaviour when the system's security is breached or there are accidental failures.
- Guaranteeing information properties under cyber-attacks and adversarial AI (quality, coherence, integrity, reliability, etc.).
- Ensuring safety, security and privacy of embedded intelligence (c.f. Chapter 1.3 "Embedded Software and beyond").
- Guaranteeing a system's coherence among different heterogeneous requirements (i.e. secure protocols, safety levels, computational level needed by the promoted

mechanisms) and different applied solutions (i.e. solutions for integrity, confidentiality, security, safety) in different phases (i.e. design, operation, maintenance, repair and recovery).

- For safety-critical applications, the open-source software (for instance: Virtual Prototypes, compilers and linkers, debuggers, programmers, integrated development environments, operating systems, software development kits and board support packages) must be qualified regarding functional safety and security standards in order to offer a possibility to create transparent, auditable processes for ensuring safety and security (c.f. Chapter 3.1 "Mobility").
- Assess complex System-on-Chip implementations and a Chiplet approach assembling functional circuit blocs with different functions (e.g.: processor, accelerator, memories, interfaces, etc.) regarding security functionalities, focusing on scalability, modularity as well as Edge paradigm.
- Developing rigorous methodology supported by evidence to prove that a system is secure and safe, thus achieving a greater level of transparency without compromising information and trustworthiness.
- Evaluating the environmental impact of the implemented safety and security solutions (the green chapter connection).

2.4.4.4. Major Challenge 4: Ensuring of safety and resilience

2.4.4.4.1 State of the art

Safety has always been a key concept at the core of human civilisation. Throughout history, its definition, as well as techniques to provide it, has evolved significantly. In the medical application domain, for example, we have witnessed a transformation from safe protocols to automatic medication machines, such as insulin pumps and respiratory automation, which have integrated safety provisions. Today, we can build a range of different high-integrity systems, such as nuclear power plants, aircraft and autonomous metro lines. The safety of such systems is essentially based on a combination of key factors, including: (i) determinism (the system's nominal behaviour is always the same under the same conditions); (ii) expertise and continuous training of involved personnel; (iii) deep understanding of nominal and degraded behaviours of the system; (iv) certification/qualification; and (v) clear liability and responsibility chains in the case of accidents.

This context has been considerably challenged by the predominant use of AI-based tools, techniques and methods. A capital example is the generative AI, which is forcefully and naturally making its way into the digitalization of ubiquitous electronic components and systems.

Techniques based on Machine Learning, generative AI and more generally AI are used mainly in two ways: embedded in ECS and as a tool for carrying out safety analysis.
Much has been written about the limits of traditional safety techniques, which need to be extended and/or embedded in new overall safety-case arguments, whenever ECS embeds IA. In comparing, much less has been said about the use of AI to perform safety analysis in compliance with the regulations.

To govern this contest, a couple of years ago Europe published AI act. The international standardization group ISO/IEC JTC 1 is hardly working and publishing a set of standards, which are character to being domain independent applications. Even in the nuclear domain, certainly among the most restrictive and conservative ones, the related industrial community is investigating, not without a live internal debate, on how to use of AI-based techniques, identify the limits and study their impacts on safety (see e.g., the activities and works of IEC TC 45 SC 45A and the International Atomic Energy Agency).

This major challenge is devoted to understand and develop innovations, which are required to increase the safety and resilience of systems in compliance with AIact and other related standards, by tackling key-focus areas involving cross-cutting considerations such as legal concerns and user abilities, and to ensure safety-related properties under a chiplet-based approach (c.f. introduction).

2.4.4.4.2 Vision and expected outcome

The vision points to the development of safe and resilient autonomous systems in dynamic environments, with a continuous chain-of-trust from the hardware level up to the applications that is involved in the accomplishment of the system's mission, including AI. Our vision takes into account physical limitations (battery capacity, quality of sensors used in the system, hardware processing power needed for autonomous navigation features, etc.), interoperability (that could be brought e.g. via open source hardware), and considers optimizing the energy usage and system resources of safety-related features to support sustainability of future systems. Civilian applications of (semi-) autonomous mobile systems are increasing significantly.

This trend represents a great opportunity for European economic growth. However, unlike traditional high-integrity systems, the hypothesis that only expert operators can manipulate the final product undermines the large-scale adoption of the new generation of autonomous systems.

Civilian applications thus inherently entail safety, and in the case of an accident or damage (for example, in uploading a piece of software in an AI system) liability should be clearly traceable, as well as the certification/qualification of AI systems.

In addition to the key focus areas below, the challenges cited in Chapter 2.3 on Architecture and Design: Methods and Tools are also highly relevant for this topic, and on Chapter 1.3 on Embedded Software and beyond.

2.4.4.4.3 Key focus areas

2.4.4.4.3.1 Dynamic adapdation and configuration, self-repair capabilities, (decentralized instrumentation and control for) resilience of complex and heterogeneous systems

The expected outcome is systems that are resilient under physical constraints and are able to dynamically adapt their behaviour in dynamic environments:

- Responding to uncertain information based on digital twin technology, run-time adaption and redeployment based on simulations and sensor fusion.
- Automatic prompt self-adaptability at low latency to dynamic and heterogeneous environments.
- Architectures, including but not limited to the RISC-V ones that support distribution, modularity and fault containment units to isolate faults, possibly with run-time component verification.
- Use of AI in the design process – e.g. using ML to learn fault injection parameters and test priorities for test execution optimization.
- Develop explainable AI models for human interaction, systems interaction and certification.
- Resource management of all systems' components to accomplish the mission system in a safe and resilient way. Consider to minimize the energy usage and system resources of safety-related features to support sustainability of future cyber-physical systems.
- Identify and address transparency and safety-related issues introduced by AI applications.
- Support for dependable dynamic configuration and adaptation/maintenance to help cope with components that appear and disappear, as ECS devices to connect/disconnect, and communication links that are established/released depending on the actual availability of network connectivity (including, for example, patching) to adapt to security countermeasures.
- Concepts for SoS integration, including legacy system integration.

2.4.4.4.3.2 Modular certification of trustable systems and liability

The expected outcome is clear traceability of liability during integration and in the case of an accident:

- Having explicit workflows for automated and continuous layered certification/qualification, both when designing the system and for checking certification/qualification during run-time or dynamic safety contracts, to ensure continuing trust in dynamic adaptive systems under uncertain and/or dynamic environments.

- Concepts and principles, such as contract-based co-design methodologies, and consistency management techniques in multi-domain collaborations for trustable integration.
- Certificates of extensive testing, new code coverage metrics (e.g. derived from mutation testing), and formal methods providing guaranteed trustworthiness.
- Ensuring trustworthy electronics, including trustworthy design IPs (e.g. source code, documentation, verification suites) developed according to auditable and certifiable development processes, which give high verification and certification assurance (safety and/or security) for these IPs.

2.4.4.4.3.3 Safety aspects related to the human/system interaction

The expected outcome is to ensure safety for the human and environment during the nominal and degraded operations in the working environment (cf. Major Challenge 5 below):

- Understanding the nominal and degraded behaviour of a system, with/without AI functionality.
- Minimising the risk of human or machine failures during the operating phases.
- Ensuring that the human can safely interface with machine in complex systems and SoS, and also that the machine can prevent unsafe operations.
- New self-learning safety methods to ensure safety system operations in complex systems.
- Ensuring safety in machine-to-machine interaction.

2.4.4.5. Major Challenge 5: Human systems integration

2.4.4.5.1 State of the art

This ECS SRIA roadmap aligns the RD&I for electronic components to societal needs and challenges. The societal benefits thereby motivate the foundational and cross-sectional technologies as well as the concrete applications in the research agenda. Thereby, many technological innovations occur on a subsystem level that are not directly linked to societal benefits themselves until assembled and arranged into larger systems. Such larger systems then most of the time require human users and beneficiary to utilize them and thereby achieve the intended societal benefits. Thereby, it is common that during the subsystem development human users and beneficiaries stay mostly invisible. Only once subsystems are assembled and put to an operational system, the interactions with a human user become apparent. At this point however, it is often too late to make substantial changes to the technological subsystems and partial or complete failure to reach market acceptance and intended societal benefits can result. To avoid such expensive and resource intensive failures, Human Systems Integration (HSI) efforts attempt to accompany technological maturation that is often measured as Technological Readiness Levels (TRL) with the maturation of Human Readiness Levels (HRL). Failures to achieve high HRL beside high TRLs have been demonstrated in various domains such as military, space travel, and aviation. Therefore, HSI efforts to achieve high HRLs need to be appropriately planned, prepared, and coordinated as part of technological innovation cycles. As this is currently only rarely done in most industrial R&D activities, this Chapter describes the HSI challenges and outlines a vision to address them.

There are three high-level HSI challenges along ECS-based products:

- The first challenge is to design products that are acceptable, trustworthy, and therefore highly likely to be used sustainably to achieve the expected individual, organizational, or societal benefits. Thereby, the overall vision for the practical use of a product by real users within their context is currently often unknown at the time of the technological specification of the product. Instead, the technological capabilities available in many current innovation environments are assembled to demonstrate merely technological capabilities but not operational use. This is often mistakenly called “use case” as it means “use to demonstrate the product” not “realistic use of the product by real users in realistic environments". Thereby, sufficiently detailed operational knowledge of the environmental, organizational, and user characteristics is often not available and cannot be integrated into technology development. Therefore, the conception of accepted and trusted, and sustainably used technologies is often the result of trial-and-error, rather than strategically planned development efforts.

- The second challenge consists of currently prevalent silos of excellence where experts work within their established domain without much motivation, ability or interest to requirements that seem external to their domain. Instead, success is often seen as promoting the own area of expertise. This forms effective resistance against a holistic design of a system instead of subcomponent optimization and makes it difficult to design products for accepted, trusted, and sustained usage. For example, increasingly complex and smart products require often intricate user interactions and understanding that is beyond simpler "non-smart" products. Thereby, the developing engineers often do not know the concrete usage conditions of the to-be-developed system or constraints of their users and are therefore unable to make appropriate architecture decisions. For example, drivers and workers generally do not like to purely monitor or supervise automated functions, while losing active participation. This is especially critical when humans have to suddenly jump back into action and take control when unexpected conditions require them to do so. Therefore, aligning the automation capabilities with the human tasks that are feasible for users to perform and to match their knowledge and expectable responsibilities, are becoming paramount to bring a product successfully to the market. However, currently established silos of engineering excellence in organizations are difficult to penetrate and therefore resist such as external perceived requirements.

Thirdly, continuous product updates and maintenance are creating dynamically changing products that can be challenging for user acceptance, trust, and sustained usage. Frequent and increasingly automated software updates have become commonplace to achieve sufficiently high security levels and to enable the latest software capability sets as well as allow self-learning algorithms to adapt to user preferences, usage history, and environmental changes. However, such changes can be confusing to users if they come unannounced, or are difficult to understand. Also, the incorrect usage that may result from this may lead to additional security and acceptance risks. Therefore, the product maintenance and update cycles need to be appropriately designed within the whole product lifecycle to ensure maximum user acceptance and include sufficient information on the side of the users. Here HSI extends beyond initial design and fielding of products.

2.4.4.5.2 Vision and expected outcome

The vision and expected outcome is that these three HSI challenges can be addressed by appropriately orchestrating the assessments of needs, constraints, and abilities of the human users, and use conditions with the design and engineering of products as well as their lifetime support phases. Specifically, this HSI vision can be formulated around three cornerstones:

- Vision cornerstone 1: conceiving systems and their missions, based on a detailed analysis of acceptance and usage criteria during the early assessment of the usage context.

This specifically entails the assessment of user needs and constraints within their context of use and the translation of this information into functional and technical requirements to effectively inform system design and development. Such information is currently not readily available to system architects, as such knowledge is currently either hidden or not assessed at the time when it is needed to make an impact during system conception. Instead, such assessments require specific efforts using the expertise of social scientists such as sociologists, psychologists, and human factors researchers who have also familiarity or training in engineering processes. As part of this cornerstone, assessments are conducted that describe the user population and the usage situation including criticality, responsibilities, environment, required tasks and time constraints. Also, the organizational conditions and processes within which the users are expected to use the system play an important role that should impact design decisions, for example to determine appropriate explainability methods. This assembled information is shaping the system architecture decisions and is formulated as use cases, scenarios, and functional and technical requirements.

- Vision cornerstone 2: to translate the foundational requirements from cornerstone 1 into an orchestrated system mission and development plan using a holistic design process. Multifaceted developer communities thereby work together to achieve acceptable, safe, and trustworthy products. Thereby, the product is not designed and developed in isolation but within actively explored contextual infrastructures to bring the development and design communities close to the use environment and conditions of the product. Considering this larger contextual field in the design of products requires advanced R&D approaches and methodologies, to pull together the various fields of expertise and allow mutual fertilization. This requires sufficiently large, multi-disciplinary research environments for active collaboration and enablement of a sufficient intermixture between experts and innovation approaches. This also requires virtual tool sets for collaboration, data sharing, and solution generation.
- Vision cornerstone 3: detailed knowledge about the user and use conditions are also pertinent to appropriately plan and design the continuous adaptations and updates of products during the lifecycle. Converging user knowledge and expectations will allow more standardized update policies. This will be addressed by bringing the end-users, workers, and operators toward achieving the digital literacy with a chance to enable the intended societal benefits. The formation of appropriate national and international training and educational curricula will work toward shaping users with sufficiently converging understanding of new technology principles and expectations as well as knowledge about responsibilities and common failure modes to facilitate sustained and positively perceived interactions.

Within these cornerstones, the vision is to intermingle the multi-disciplinary areas of knowledge, expertise, and capabilities within sufficiently inter-disciplinary research and development environments where experts can interact with stakeholders to jointly design, implement, and test novel products. Sufficiently integrated simulation and modeling that includes human behavioral representations are established to link the various tasks. The intermingling starts with user needs and contextual assessments that are documented and formalized sufficiently to stay available during the development process. Specifically, the skills and competences are formally recorded and made available for requirements generation.

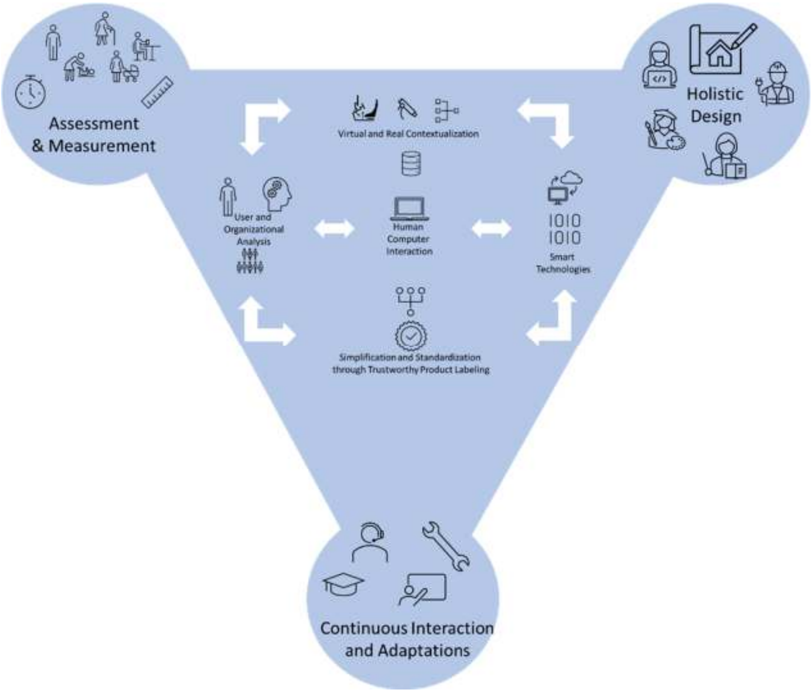


Figure 2.4.3 - Human Systems Integration in the ECS SRIA

2.4.4.5.3 Key focus areas

- Systematize methods for user, context, and environment assessments and sharing of information for user-requirement generation. Such methods are necessary to allow user centered methods to achieve an impact on overall product design.
- Develop simulation and modeling methods for the early integration of Humans and Technologies. The virtual methods link early assessments, holistic design activities, and lifelong product updates and bring facilitate convergence among researchers, developers, and stakeholders.
- Establish multi-disciplinary research and development centers and sandboxes. Interdisciplinary research and development centers allow for the intermingling of experts and stakeholders for cross-domain coordinated products and life-long product support.

2.4.5. TIMELINE

MAJOR CHALLENGE	TOPIC	SHORT TERM (2025-2029)	MEDIUM TERM (2030–2034)	LONG TERM (2035 and beyond)
Major Challenge 1: Ensuring HW quality and reliability	**Topic 1.1:** quality: *in situ* and real-time assessments	• Create an environment to fully exploit the potential of data science to improve efficiency of production through smart monitoring to facilitate the quality of ECS and reduce early failure rates	• Establish a procedure to improve future generation of ECS based on products that are currently in the production and field → feedback loop from the field to design and development	• Provide a platform that allows for data exchange within the supply chain while maintaining IP rights
	Topic 1.2: reliability: tests and modelling	• Development of methods and tools to enable third generation of reliability – from device to SoS	• Implementation of a novel monitoring concept that will empower reliability monitoring of ECS	• Identification of the 80% of all field-relevant failure modes and mechanisms for the ECS used in autonomous systems
	Topic 1.3: design for (EoL) reliability: virtual reliability assessment prior to the fabrication of physical HW	• Continuous improvement of EDA tools, standardisation of data exchange formats and simulation procedures to enable transfer models and results along full supply chain	• Digital twin as a major enabler for monitoring of degradation of ECS	• AI/ML techniques will be a major driver of model-based engineering and the main contributor to shortening the development cycle of robust ECS
	Topic 1.4: PHM of ECS: increase in functional safety and system availability	• Condition monitoring will allow for identification of failure indicators for main failure modes	• Hybrid PHM approach, including data science as a new potential tool in reliability engineering, based on which we will know the state of ECS under field loading conditions	• Standardisatio n of PHM approach along all supply chains for distributed data collection and decision-making based on individual ECS
Major Challenge 2: Ensuring dependability in connected software	**Topic 2.1:** dependable connected software architectures	• Development of necessary foundations for the implementation of dependable connected software to be extendable for common SW systems (open source, middleware, protocols)	• Set of defined and standardised protocols, mechanisms and user-feedback methods for dependable operation	• Widely applied in European industry
	Topic 2.2: dependable softwarisation and virtualisation technologies	• Create the basis for the increased use of commodity hardware in critical applications	• Definition of softwarisation and virtualisation standards, not only in networking but in other applications such as automation and transport	• Efficient test strategies for combined SW/HW performance of connected products

	Topic 2.3: combined SW/HW test strategies	• Establish SW design characteristics that consider HW failure modes	• Establish techniques that combine SW reliability metrics with HW reliability metrics	•
Major Challenge 3: Ensuring privacy and cybersecurity	**Topic 3.1:** trustworthiness	• Root of trust system, and unique identification enabling security without interruption from the hardware level up to applications, including AI	• Definition of a framework providing guidelines, good practices and standards oriented to trust	• Developing rigorous methodology supported by evidence to prove that a system is secure and safe, thus achieving a greater level of trustworthiness
	Topic 3.2: security and privacy- by-design	• Establishing a secure and privacy-by-design European data strategy and data sovereignty	• Ensuring the protection of personal data against potential cyber-attacks in the data-driven digital economy • Ensuring performance and AI development (which needs considerable data) by guaranteeing GDPR compliance	• Provide a platform that allows for data exchange within the supply chain while maintaining IP rights
	Topic 3.3: ensuring both safety and security properties	• Guaranteeing information properties under cyber-attacks (quality, coherence, integrity, reliability, etc.) independence, geographic distribution, emergent behaviour and evolutionary development	• Ensuring the nominal and degraded behaviour of a system when the underlying system security is breached or there are accidental failures • Evaluating the impact of the contextualisation environment on the system's required levels of safety and security	• Identification of the 80% of all field-relevant failure modes and mechanisms for the ECS used in autonomous systems
Major Challenge 4: Ensuring safety and resilience	**Topic 4.1:** safety and resilience of (autonomous AI) systems in dynamic environments	• Resources' management of all system's components to accomplish the mission system in a safe and resilient way • Use of AI in the design process – e.g. using ML to learn fault injection parameters and test priorities for test execution optimization	• Apply methods for user context and environment assessments and sharing of information for stakeholder-requirement generation to prototypical use cases, establish practices of use and generally applicable tools	• Develop standard processes for stakeholder context and environment assessments and sharing of information • Develop standard processes for stakeholder knowledge, skills, and competence capturing techniques to inform requirements generation • Develop educational programs to increase the levels of common stakeholder knowledge, skills and competences for sustainable product uptake across Europe
	Topic 4.2: modular certification of trustable systems and liability	• Contract-based co-design methodologies, consistency management techniques in multi-domain collaborations	• Definition of a strategy for (modular) certification under uncertain and dynamically changing environments • Consolidation of a framework providing guidelines, good practices and standards oriented to trust • Ensuring compliance with the AI standards	• Ensuring liability

	Topic 4.3: dynamic adaptation and configuration, self-repair capabilities (decentralised instrumentation and control for), resilience of complex systems	• Support for dependable dynamic configuration and adaptation/ maintenance • Concepts for SoS integration, including the issue of legacy system integration • Using fault injection methods, models-of-the-physics and self-diagnostic architecture principles to understand the true nature of the world, and respond to uncertain information (included sensor's false positives) or attacks in a digital twin, run-time adaptation and redeployment based on simulations and sensor fusion • Architectures that support distribution, modularity and fault containment units to isolate faults, possibly with run-time component verification	• Guaranteeing a system's coherence while considering different requirements, different applied solutions, in different phases	•
	Topic 4.4: safety aspects related to HCI	• Minimising the risk of human or machine failures during the operating phases • Ensuring that the human can safely interface with the machine, and also that the machine prevents unsafe operations • Ensuring safety in machine-to-machine interaction	• Develop prototypical use cases where interdisciplinary research and development centers allow for the intermingling of experts and stakeholders for cross-domain coordinated products and life-long product support.	•

Major Challenge 5: Human–systems integration	**Topic 5.1**: Establish skills and competences needed for engineering and management to jointly perform user, context, and environment assessments for user-requirement requirements generation	• Establish research lighthouses for HSI by establishing examples for effective HSI during product design, development and operation. • Investigate through research the necessary individual knowledge, skills and common practices for effective HSI integration, on individual, process, and organizational level. • Establish stakeholder knowledge, skills, and competence capturing techniques to inform requirements generation	• Bring the results of the short term activities on Topic 5.1 toward policy recommendations for education, development, and practice. • Based on the short term activities on topic 5.1, develop recommendations for appropriate education to promote HSI for socio-technical developments and operations • Based on the short term activities on topic 5.1, develop recommendations for appropriate tools and processes to promote HSI for socio-technical developments and operations • Based on the short term activities on topic 5.1, develop recommendations on organizational prerequisites to promote HSI for socio-technical developments and operations	• Based on the medium term recommendations on topic 5.1, develop policies and standards, as well as sponsoring funding schemes for training and educational programs that facilitate HSI in socio-technical developments and operations. • Based on the medium term recommendations on topic 5.1, develop policies and standards, as well as sponsoring excellence and standardization centers to establish common and standardized tools and virtual methods that facilitate HSI in socio-technical developments and operations. • Based on the medium term recommendations on topic 5.1, develop policies and standards for organization certifications of HSI in socio-technical developments and operations.
	Topic 5.2: Develop simulation and modeling methods for the early integration of Humans and Technologies	• Create tools that allow to link early assessments, holistic design activities, and lifelong product updates to facilitate convergence among researchers, developers, and stakeholders communities • Establish tools to bring stakeholder knowledge, skills, and competence capturing techniques to inform design and development activities • Establish tools to quantify risks of human acceptance and trust • Establish tools to collect and share data bases on relevant human behavioral metrics (safety, acceptance, trust)	• To establish and promote the tools and methods identified during the short term activities for topic 5.2, establish centers of excellence for HSI for socio-technical systems, focusing on promoting early user need and constraint assessments, holistic design activities, and lifelong product updates. The centers of excellence should be harmonized internationally but reflect the idiosyncrasies of individual member states and situations. From the mid-term on, topic 5.2 and topic 5.3 are merged.	• Establish holistic design and systemic thinking education and training in technical and social-sciences academic and non-academic educational programs across Europe and individual member states to promote the knowledge and experience gained the centers of excellence.

	Topic 5.3: establish multi-disciplinary research and development centers and sandboxes	• Establish interdisciplinary research and development centers allow for the intermingling of experts and stakeholders for cross-domain coordinated products and life-long product support. • Establish tools and processes to update stakeholder knowledge, skills, and competence capturing techniques to inform design and development activities	• To establish and promote the tools and methods identified during the short term activities for topic 5.2, establish centers of excellence for HSI for socio-technical systems, focusing on promoting early user need and constraint assessments, holistic design activities, and lifelong product updates. The centers of excellence should be harmonized internationally but reflect the idiosyncrasies of individual member states and situations. From the mid-term on, topic 5.2 and topic 5.3 are merged.	• Establish holistic design and systemic thinking education and training in technical and social-sciences academic and non-academic educational programs across Europe and individual member states to promote the knowledge and experience gained the centers of excellence.

3

Strategic Research and Innovation Agenda 2025

ECS KEY APPLICATION AREAS

3.1

ECS Key Application Areas

MOBILITY

3.1 Mobility

3.1.1 Scope

Mobility is a basic human need and Europe's mobility industry is a key contributor to it. The automotive sector alone provides employment, both direct and indirect, to 13.8 million Europeans, representing 6.1% of total EU employment. 2.6 million people work in the direct manufacturing of motor vehicles, representing 8.5% of EU employment in manufacturing[1]. The automotive sector is also the driver for innovation in many other mobility sectors in Europe, including aerospace, maritime and rail, but also in other sectors as e.g. farming and industry (advanced robotics).

The JRC Report "The future of road transport[2]" from April 2019 states: "Four key game changers are shaping the future of road transport: automation, connectivity, decarbonisation and sharing. These future technologies and services promise to contribute to fewer negative impacts from road transport while also generating new mobility paradigms and transport governance opportunities."

This results in two major technological transformations for the mobility sector: the decarbonisation, which is the transition to a CO_2-neutral mobility as contribution to the Green Deal, and the digitalization, which is in essence the introduction of significantly more automation in mobility to cope with societal challenges in an aging society without deviating from the vision of zero fatalities, as well as to provide mobility for a growing global population faced with scarce resources supported by environmentally optimal mobility modes and a CO_2-neutral mobility.

Therefore, the ECS-SRIA mobility chapter part dealing with CO2 free mobility is aligned with the proposal for the partnership "Towards zero emission road transport" (2Zero) programme by Horizon Europe to achieve carbon-neutrality in road transport by 2050. Cooperation between 2Zero and the ECS community is important and well established.

This requires research in automated driving and ADAS advancing automotive safety, efficiency, and convenience. Relevant key focus areas include sensor fusion, control algorithms, computer vision, and real-time testing and verification. As the demand for ADAS and AD grows, understanding consumer preferences and addressing the technical challenges will be essential. Additionally, given the potential overlap between ADAS and fully autonomous vehicles, research efforts must align with broader industry trends and explore how existing technologies can pave the way for self-driving cars. Therefore, this mobility chapter is closely aligned with the activities of the partnership "Connected, Cooperative and Automated Mobility" (CCAM) under Horizon Europe.

These two major trends lead towards software defined vehicles, where customers replace horsepower by software functionality in their decision process for vehicles and mobility services (or modes). The Vision and Roadmap document[3] of European Software-Defined Vehicle of the Future (SDVoF) initiative states: *Software now drives value creation, serving functions and services both within vehicles (on-board) and in the cloud (off-board) as well as the infrastructure around the vehicle, which will provide mobility services. Customers prioritize "software freshness," seeking new applications and services related to infotainment,*

[1] Internal Market, Industry, Entrepreneurship and SMEs, https://ec.europa.eu/growth/ sectors/automotive_en
[2] https://publications.jrc.ec.europa.eu/repository/handle/JRC116644, JRC publication repository, Apr 2019
[3] https://federate-sdv.eu/wp-content/uploads/2024/04/2024-04-12-SDVoF-Vision-document-ver017-final.pdf, April 2024

connectivity, and ADAS/AD functionality. Regular over-the-air updates enhance cyber-security, safety, and innovation during the vehicle's operational lifespan. This transition fuels demand for next-generation system-on-chip designs and high-performance processors, fundamentally reshaping software development and integration, and opens the opportunity to re-think and re-design the vehicle software stack to match the need of the vehicle of the future.

Regaining European sovereignty in semiconductors for mobility is one important goal in these transitions. The other one is to avoid disconnection between cities with abundant infrastructure and less developed regions with lack of understanding for new technologies, but need for basic mobility and services. Europe is ranked number one in automotive semiconductors. In the automotive value chain, Tier 1's and original equipment manufacturers (OEMs) are also top global players and intend to gain further market share through close collaboration with semiconductor and embedded software leaders in Europe.

In general, the ECS community needs to think also politically and take into account the fact that some parts of the population feel lost in all these transitions. Step back, slow down, analyse what might help immediately, do development works in the regions rather than in cities.

Make sure that developments are not only done for initial prototyping and further departure to Asia.

The electronics components and systems (ECS) community will contribute substantially to these tasks by using new technologies, components and systems to target the following topics:

- Electrification of vehicles and development of powertrains for carbon-free energy carriers. Enabling technologies are coming from the European ECS industry – for instance, energy-efficient devices, power electronic components and systems, energy (e.g. battery) management systems, and embedded software solutions for power management. Furthermore, integration into the grid to enable better contribution to the energy transition with bidirectional energy flow. Strengthening of infrastructure to increase confidence into electric mobility and exploitation of opportunities to demonstrate cost advantages of EV compared to ICE cars. Strengthen European technologies through fostering collaboration between actors in vertical value chains.

- Automated and autonomous vehicles and coordinated mobility to make traffic more efficient and thus reduce pollution by new electronics architectures, smart and connected sensor systems, AI-based real-time software, higher performance in-vehicle controllers and networks, as well as connectivity devices and advanced embedded software solutions. Implementation of mobility services to close gaps for users in regions with limited mobility infrastructure.

- Rapid advances in AI and edge computing will ensure Europe can produce a step change in these areas. Automated and autonomous driving, mobility and logistics are complex applications where the use of AI technologies is growing very rapidly, affecting both society and industry directly. The European transport industry is being revolutionised by the introduction of AI (combined with electric vehicles). However, AI applications in transport are very challenging, as they typically have to operate in highly complex environments, a large number of possible situations and real-time, safety- relevant decision-making. Leading IT companies in the US and China in particular are providing a challenge to European industry in these areas, and significant effort will be required to safeguard the leading position of the European automotive industry. Hence the development of high-definition digital maps, digital twins is a crucial element to further develop and update all types of vehicles interacting with and within the environment, with each other and with the cloud. It must be accompanied by evolutionary tools and processes that allow employees to learn, adapt and evolve into new assignments and tasks.

- There will be a need to more intensively monitor, interact with, and update the car remotely. Remote supervision and operation as well as software upgrades need to be done over-the-air (OTA) via high-speed connectivity. Semiconductors have a shorter lifetime than cars and hence it is needed to predict their end-of-life and to replace them on-time before the car breaks down.

Mobility products are used by people/citizens, everywhere and every time. It means that it has **impact**, that has to be approached with different perspectives:

- Human: equity, inclusivity, health
- Safety, (cyber)security
- Accessibility (incl. affordability)
- Environment

Each perspective is a powerful innovation driver, requiring ECS novel solutions such as

- Human: driving under influence (drugs, alcohol…), emergency calls, devices for disabled people, health monitoring…
- Safety: child presence detection, Vulnerable Road Users…
- Affordability: ECS guidelines to achieve affordable cars?
- Environment: eco-design, circularity …

This impact also leads to new regulations and policies on EC or national level which need to be considered already at very early stages.

The global automotive industry is already undergoing this transformative shift due to software-defined vehicle (SDV) concepts and over-the-air (OTA) updates. SDVs separate hardware from software, allowing for dynamic updates and upgrades. The vehicle becomes a command center, collecting and organizing vast data volumes, applying AI insights, and automating actions. This architecture transformation simplifies integration and supports service-based business models. Meanwhile, OTA updates enable vehicles to add new features, enhance performance, and address issues without physical visits to service centers. Consumers experience progressively improves during vehicle ownership. Customized enhancements and safety feature updates create a closer relationship between original equipment manufacturers (OEMs) and end-users, opening also new opportunities to developers. As a result, the automotive landscape is becoming more connected, adaptable, and user-centric, with significant implications for business models.

The emerging mobility applications in SDV typically comprise both hardware (HW) and software (SW) components within the vehicle, as well as additional components hosted in the cloud leading to edge-to-cloud (edge2cloud) applications. These cloud-based elements rely on data from various industry sectors demanding a coordinated approach in their respective development roadmaps. For instance, optimised charging applications guide electric vehicles to available and fast charging stations along their routes. These applications rely on real-time information about vehicles battery status and the charging station status, the integration with wireless billing systems, and -in the case of bidirectional charging - close ties to energy providers. Other examples include automatic toll collection systems, parking and depot applications, and traffic control systems that provide correct road condition information (including hazards like black ice) based on data shared by other vehicles, weather authorities and traffic/transport infrastructure. The complexity of these edge2cloud applications increases significantly due to the need for coordinated development across diverse industry domains.

In addition, the European automotive industry faces intensified global competition due to non-EU manufacturers' early adoption of software-driven strategies. Large tech companies and hyper-scalers leveraging substantial software budgets and indirect business models, are already dominating specific domains. Additionally, significant state aid in East Asia facilitates rapid market entry for new companies.

Key is an user-oriented innovation: the major competitiveness driver for an application industry such as automotive is to quickly introduce user-oriented innovations, i.e., innovations that bring maximum "visible" value to the customer. It means that the main challenge is to integrate the best ECS technologies into innovative use-cases. Electronics- and software-based systems provide important key enabling technologies for these transformations. Sensors, semiconductors and embedded AI enabled software in integrated intelligent systems are essential building blocks. Nevertheless, energy and cost efficiency in the operation of transport systems as well as in the development and production of mobility systems as vehicles, ships, trains, airplanes, remain important RDI tasks.

Therefore, the research scope encompasses the development of energy efficient mobility grade semiconductor, sensors, actuators, AI enabled HW platforms as well as embedded AI enabled software platforms and edge2cloud software for all mobility modes. Automotive mobility includes systems for passengers and goods (passenger cars, two/three wheelers, trucks), mobility in smart farming with specialized farming machinery, ships and vessels in maritime mobility as well as mobility in the air (for passengers and goods) using airplanes, helicopters or UAVs (Unmanned Aerial Vehicles) (drones), and last but not least mobility on rails provided by trains.

3.1.2 Major Challenges

The changes in mobility explained above lead to following six main ECS RDI challenges for mobility:

- **Major challenge 1:** SDV hardware platforms: modular, scalable, flexible, safe & secure
- **Major challenge 2**: SW platforms for SDV of the future: modular, scalable, re-usable, flexible, safe & secure, supporting edge2cloud applications
- **Major challenge 3:** Green deal: enable climate and energy optimal mobility

- **Major challenge 4:** Digitalisation: affordable and safe automated and connected mobility for passengers and freight
- **Major challenge 5:** Edge2cloud mobility applications: added end-user value in mobility
- **Major challenge 6:** AI enabled engineering tool chain: agile collaborative SDV SW development and SDV as well as ADAS/AD validation

Mobility ECS systems consist of Hardware, Software and AI, the challenges are covering them as summarized below.

The mobility ECS system hardware comprises:

- Computing hardware as enabler for SW and SDV platforms – see challenge 1
- Hardware components for electric energy (power) and actuation – see challenge 3
- Hardware components for sensing and perception – see challenge 4

The mobility ECS system software uses information from sensing and connectivity to actuate and to manage data and interact with the driver and passengers of a vehicle (this includes the increasingly important infotainment and cockpit) – see challenge 2

AI is applied to the vehicle to improve performance, user experience, safety, etc. (see challenges 1, 2, 5) as well as applied to tooling to improve engineering & manufacturing processes (see challenge 6)

3.1.2.1 *Major Challenge 1: SDV hardware platforms: modular, scalable, flexible, safe & secure*

The automotive industry has traditionally been accustomed to a multitude of interconnected *Electronic Control Units* (ECUs), each with their own specific function, architecture, software stacks and operating systems. As more functionalities and safety features were incorporated into the vehicle, some organisation of ECUs has taken place, and different units communicate via a centralised gateway. With the evolution of the car and the functions now expected from the end-costumer, such as better infotainment services and autonomy, this model is becoming increasingly untenable. The presently prevalent decentralised architectures have significant drawbacks when it comes to scalability and communications performance.

The current trend is such that an increasing number of functionalities are software defined. In fact, it is projected that the number of lines of code per vehicle will go up from the current 100 million to 1 billion by 2030. From a hardware perspective, the increased autonomy being incorporated in the automotive sector causes that on-board computing is centralised, otherwise it would be untenable to network the increasingly interdependent ECUs together due to latency and the scale of the wiring harness, amongst other limitations.

The solution to these issues is the centralisation of the E/E architecture whereby several ECUs are consolidated into so called *Domain Control Units* (DCUs) or *Zonal Control Units (ZCUs)* that integrate different functions into a more cost-effective solution. It is foreseen that eventually these DCUs/ZCUs will be networked to a centralised SoC that integrates and combines the different functions such as autonomous driving, infotainment and cabin control.

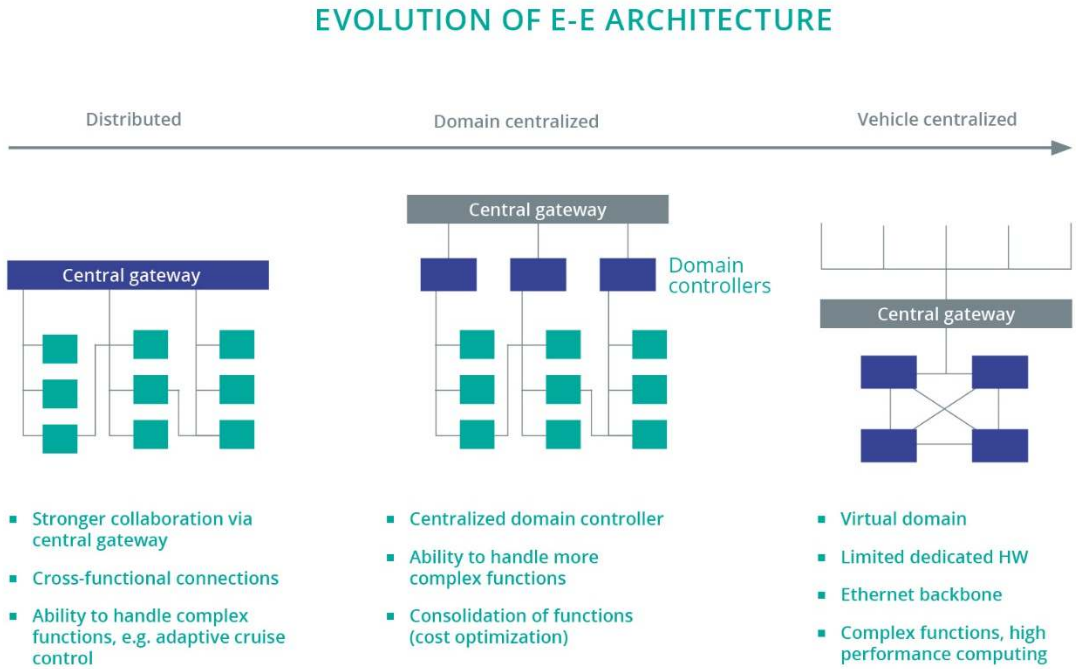


Figure 3.1-1 Evolution of E/E architecture (Source: McKinsey & Company)

With this in mind, the RISC-V Automotive Hardware Platform initiative is designed to consolidate and strengthen Europe's leadership position in the automotive electronics sector. This initiative comes at a critical time as the automotive industry faces fundamental shifts with the centralisation of electrical/electronic (E/E) architecture in vehicles and the advent of more electrified, automated, connected, and ultimately shared modes of transportation.

Furthermore, it is expected that there will be no one vehicle architecture, but that there will be different flavours, based on the type of OEM and/or based on the type of car (low-end versus high-end), as depicted in the picture below.

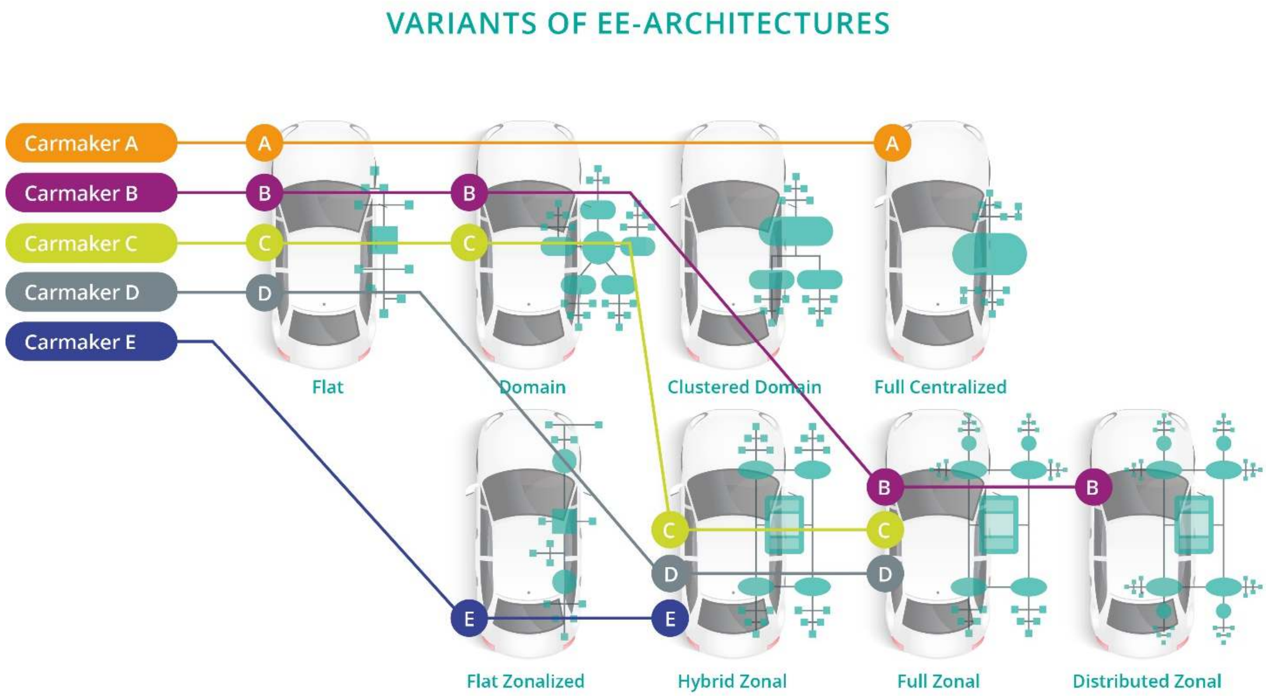


Figure 3.1-2 Variants of EE-architectures

In addition, the growing complexity in both automotive software as EE architectures (- as qualitatively shown in the picture below -) requires new families of high-end control units and application-specific processors with a large number of system integration challenges.

COMPLEXITY GROWTH IN EE ARCHITECTURES

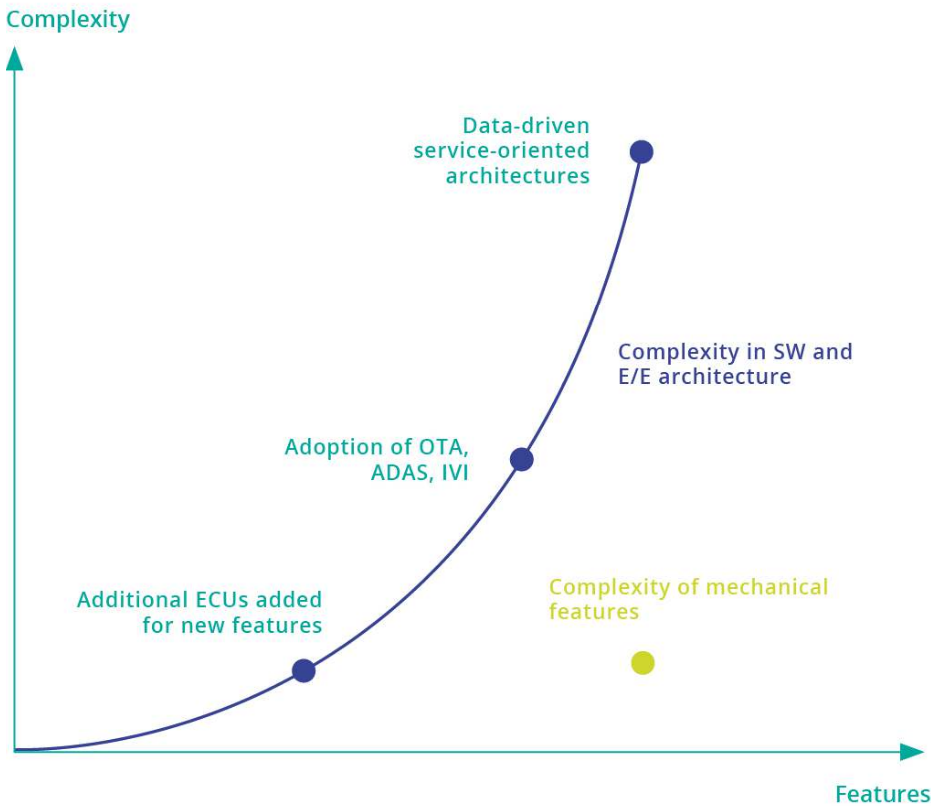


Figure 3.1-3 Complexity growth in EE architectures

To address this ongoing transition and challenge, a high-level roadmap, based on the RISC-V instruction set architecture (ISA), has been developed for designing future automotive reference platforms, such that all involved stakeholders can use the flexibility provided by the open-source ISA. Through the reference platform approach proposed by this roadmap, European industry will have a springboard for further innovation in automotive electronics.

A high-level view of the roadmap as is explained in more detail in the document[4], is given below. The roadmap has been split into 4 key elements or categories of the overall RISC-V based processors:

- Scalable RISC-V Automotive Control Processors.
- High-Performance RISC-V based Application Processors.
- AI and ML Accelerators.
- System integration and interfacing.

[4] https://ecssria.eu/ecssria-2024-downloads

– Fiber optics networks in cars.
– Wireless communication in cars incl. plug and play SW.

The four elements are linked to each other. The first contains automotive-specific real-time control processors. They are the starting point for the second category, which is more generic and applicable for other type of application domains (communications, industrial, health, etc.). One example is industrial robotics – for advanced motion control, computer vision, inspection, and health and safety applications – and drones/UAVs, where a series of grand challenges is emerging including sensor fusion and processing for detect and avoid applications in a low-cost, low-weight, low-power form factor. The third category contains accelerator elements which are specific developments for both automotive as well as other domains, and which contains starting points for next generation processors. As these accelerators are a key component of both processor families, we have opted to make a separate category to show the different IP elements to develop in the coming years. The fourth category contains specific elements for system integration, more in particular how controllers and processors fit into global system solutions.

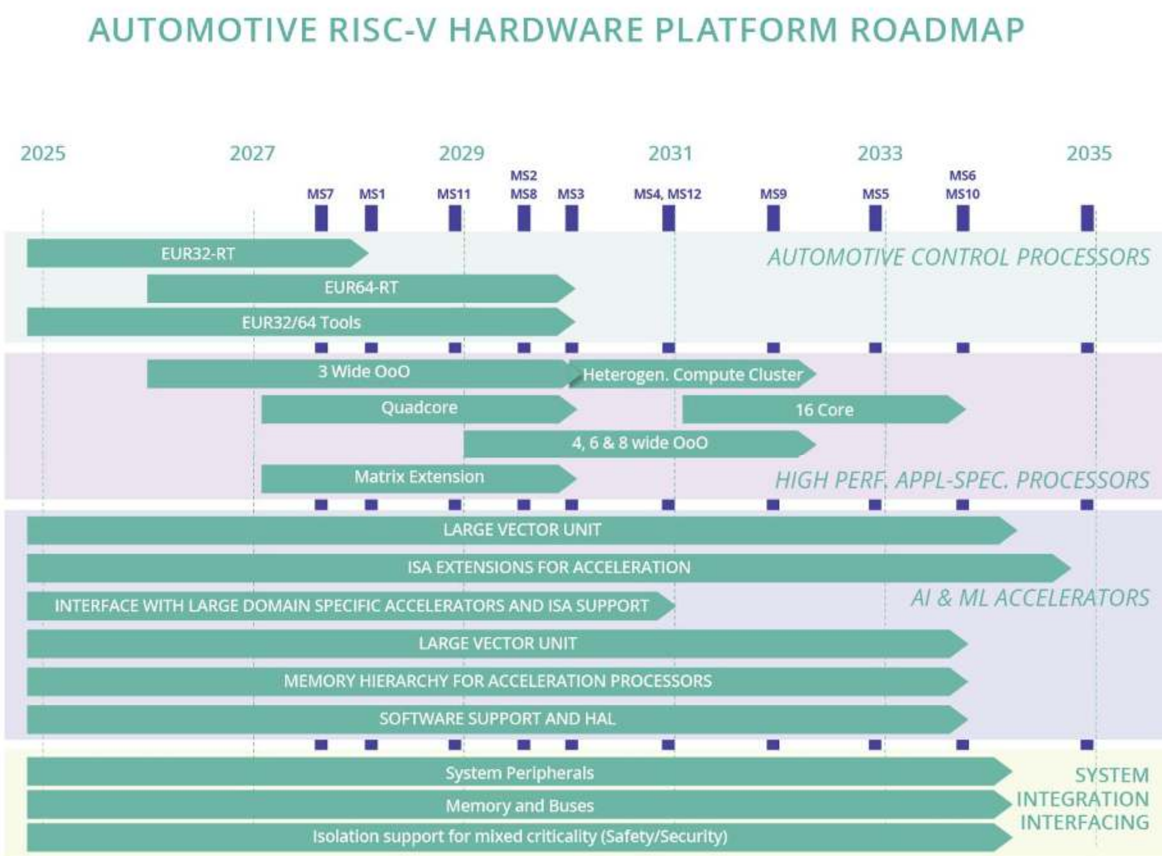


Figure 3.1-4 Automotive RISC-V hardware platform roadmap

The elements can be considered as linked activities as part of a reference platform. By definition, this includes the computational elements along with the supporting infrastructure required to realise the whole platform. Therefore, the reference platform infrastructure roadmap laid out includes both SoC and chiplets interconnects.

Along with the core (CPU) clusters, three key components are needed to sustain the high memory bandwidth needed by the processor, as well as controllability features towards supporting, to a sufficient extent, freedom from interference across real-time applications with integrity requirements. Those components are the following: (1) Shared cache memories capable of supporting a large number of outstanding requests from a high number of out-of-order cores, both in terms of (data) bandwidth as well as in terms of managing pending requests minimizing serialization, (2) High-bandwidth interconnects with Quality of Service (QoS) minimizing contention and allowing some control on the traffic priorities through the deployment of programmable QoS support, And (3) high-bandwidth memory controllers with QoS support able to manage out-of-order requests, minimizing serialization, and allowing some controllability on the traffic priorities.

Another important HW-challenge is the introduction of “automotive chiplets, which will be used in System on Chips designed considering the mobility application industry perspective and pooling for smart packaging solutions at early stage of development:

- SoC design to achieve the best trade-off between performance, integration ability and system affordability
- Explore architectures mixing advanced cutting-edge technologies as well as on-the-shelf mature solutions
- Identify key bottlenecks to be overcome
- Investigate micro-architecture solutions and set-up the playground for a fair and relevant split between “open-source” and “proprietary” zones

The development of these chiplets is part of the automotive RISC-V platform roadmap.

3.1.2.2 Major Challenge 2: SW platforms for SDV of the future: modular, scalable, re-usable, flexible, safe & secure, supporting edge2cloud applications

As the automotive industry moves towards autonomous, electric, connected, and service-oriented vehicles, hardware and software are becoming increasingly important in managing their operations and enabling new features. “Software-defined vehicles” are getting more valuable than traditional vehicles based mainly on mechanical parts, with electronics and software playing a key role in this new paradigm. Customers value new software applications such as infotainment, connectivity, ADAS/AD functionality, and regular over-the-air updates for new or improved functionality during the operational phase of the vehicles, automatically or on-demand. New apps are also combining cloud with vehicle functionalities to increase the comfort and safety of the driver for day-to-day operations such as charging, parking, and driving. Customers are already willing now to switch brands for these better applications and features.

These updates of the vehicle software over the lifetime of the vehicle results in a radical change of business models in the automotive industry. Hardware and software of vehicles are getting different lifecycles. The electrical and mechanical parts of vehicles are mostly fixed at the time when a customer buys a vehicle (- in rare cases also for HW parts replacements may be offered -), but the software continues to change, adapt and improve over the operational phase of the vehicle as long as the customer accepts software updates. The over-the-air software updates imply a significant increase in the value of a vehicle for customers. But it also requires a dedicated SW/HW abstraction in software defined vehicles.

A software platform, which includes virtualization, operating systems, middleware, and an API framework, also supporting the integration with the cloud in edge2cloud applications, plays a key role in this new paradigm of SDV. By raising attention to software and hardware, manufacturers can create more value for their customers and stay ahead of the competition. Embedded computing hardware and software are therefore have become extremely important for the OEMs.

Consequently, electronic architectures of vehicles are becoming more centralized, fuelling the demand for next-generation system-on-chip designs and high-performance processors, and redefining how software is designed, integrated, and maintained (see major challenge 1 in chapter 3.1.2.1). The software layers between hardware and applications, including interfacing with the cloud, play a key role in this paradigm shift.

Automotive players must transform themselves into software-defined companies, but they are facing difficulties with software development. Software complexity is rising sharply. Increased complexity of functionalities and sharing of computing resources across electronic control units, vehicle domains, and the mobility and cloud infrastructures reduces the software development productivity (see Figure 3.1-5). Because of the rapid transformation to software-defined vehicles (driven by new vehicle functions, features, properties), the automotive industry is facing a widening and unsustainable gap between software complexity and productivity.

Many non-compatible SW platforms (often coming from suppliers) used at different OEMs (and often even within one OEM) create big redundant and non-value adding effort in development and even more in maintenance. This leads to delays and cost overruns for software projects.

Additionally, the industry is facing a major software talent shortage.

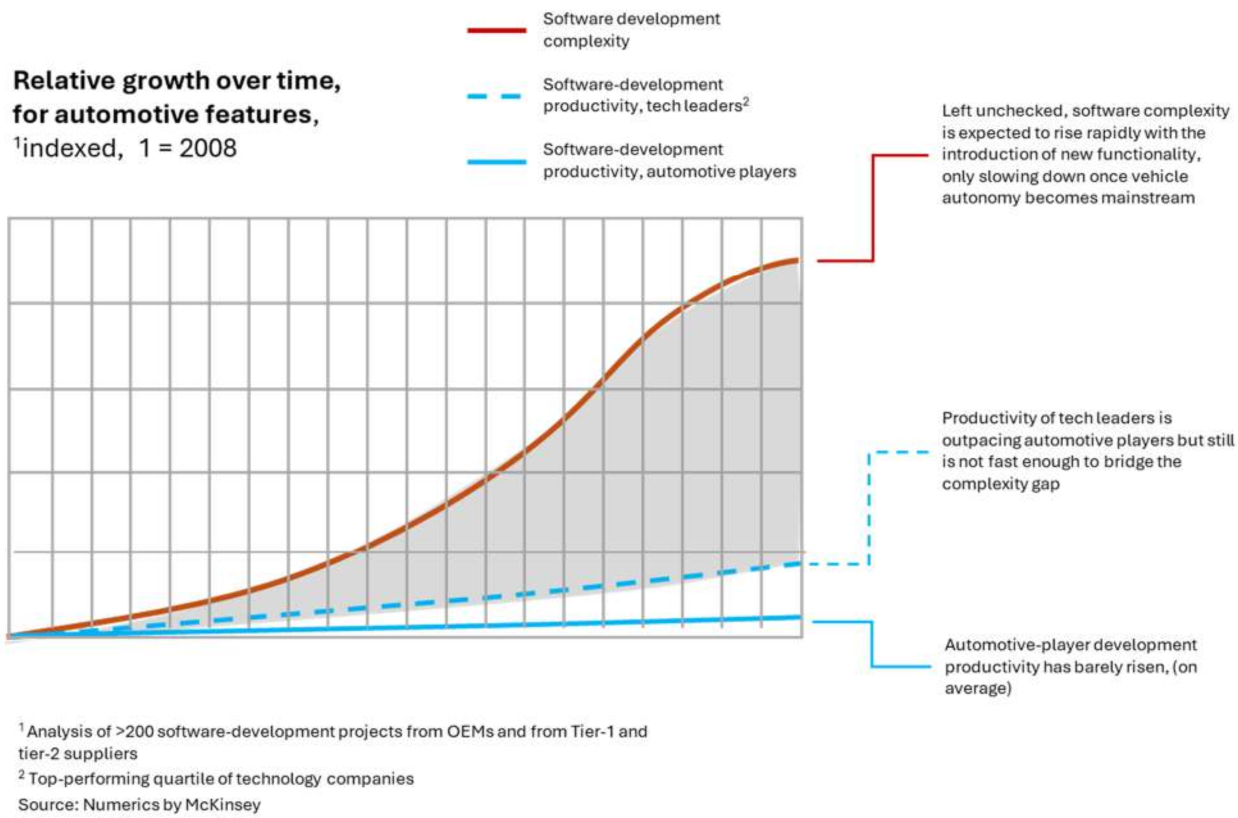


Figure 3.1-5 Increasing gap between software complexity and productivity (Source: McKinsey).

To succeed in this dynamically changing environment and to be globally competitive, companies need to minimise (or at least limit) the complexity by reducing the effort required to develop and maintain software. Consequently, the current software operating model needs to be revisited in terms of:

- Architecture, design, requirements.
- Development methodologies (e.g. agile-at-scale, or fundamental changes in development and software testing).
- Software performance management, toolchain infrastructure.
- Location and organizational unit where the software is developed, including involved partnerships.
- An end-to-end software platform (consisting of virtualisation, operating system and a middleware and API-framework layer with standardised interfaces) abstracts the hardware layer for the application. The function software layer shall be supported in managing the rising software complexity and developing future vehicles in an effective and efficient manner. Software should be truly integrated end to end, software modules should be developed on a common code base, and a primary robust operating system should be able to cover all major systems throughout the vehicle families.

So far, car companies have focused on developing proprietary technology platforms, impeding efficiencies when such investments replicate efforts on elements that are not differentiating and visible to the customer. A rising number of partnerships and alliances across varying types of actors of the automotive and digital ecosystems shows a growing openness to join forces. They however do not cover systematically all the non-differentiating elements of the software stack and lack in many cases sufficient implementation. They receive help from stronger cross-initiative coordination and governance.

In this context, the European automotive, embedded SW and semiconductor industry together with the European Commission have started complementary but distinct industry driven initiatives to reinforce EU strategic autonomy and leadership in the automotive value chain on the vehicle of the future (SDVoF). It addresses the need for an open automotive hardware platform and an open SDVoF ecosystem driven by European actors. The open SDVoF initiative focuses on an open and pre-competitive collaboration driven by European OEMs and suppliers on non-differentiating elements of the vehicle software stack. The initiative aims to reinforce the coordination between existing alliances by orchestrating distributed developments and ensuring close links with EU and global initiatives on an open automotive hardware and software platform. Stronger cross-initiative coordination and governance create benefits to existing partnerships and alliances working on these technologies.

The creation of an European SDV ecosystem is essential with the objective of developing application-oriented, open-source building blocks and to promote EU start-ups and SMEs introducing relevant mobility applications thanks to a shared toolchain. Agility and development speed are crucial challenges in this endeavor. It is mandatory to reuse existing proven building blocks and design patterns wherever possible. Cooperation with well accepted initiatives in the automotive industry working on various aspects of the SDV are essential to avoid duplication of effort and to reduce the time-to-market. Several initiatives and alliances working on an automotive SDV platform (e,.g. **AUTOSAR**[5]**, Catena**-X[6], **COVESA**[7], **Eclipse SDV**[8]**, FEDERATE**[9]**, SOAFEE**[10]**, 2Zero**[11]**, CCAM**[12]).

The initiative has agreed on a three-layer structure for SDVoF HW/SW stacks as depicted in Figure 3.1-6. The initiative aims to develop collaboratively non-differentiating building blocks in the “Middleware and API framework” and “Car meta operating system & HW abstraction” layers with implementations in-vehicle as well as off-board in the cloud. Wherever useful and possible, open-source shall help to gain agility as well as a selection of the best SW technologies and/or solutions. The building blocks will be used in existing or new SW stacks of OEMs or tiers, which will bring the new functionalities into vehicles on the

[5] Home AUTOSAR: https://www.autosar.org
[6] Catena-X: https://catena-x.net/en/
[7] COVESA: https://covesa.global/
[8] Software Defined Vehicle | The Eclipse Foundation: https://sdv.eclipse.org/
[9] FEDERATE: https://federate-sdv.eu/
[10] Soafee: https://www.soafee.io/
[11] 2Zero: https://www.2zeroemission.eu/
[12] CCAM - European Partnership on Connected, Cooperative and Automated Mobility: https://www.ccam.eu

road. Experts from the European automotive industry organized several workshops to create a layered model (see Figure 1) of SDV SW stacks as shown in Figure 3.1-6.

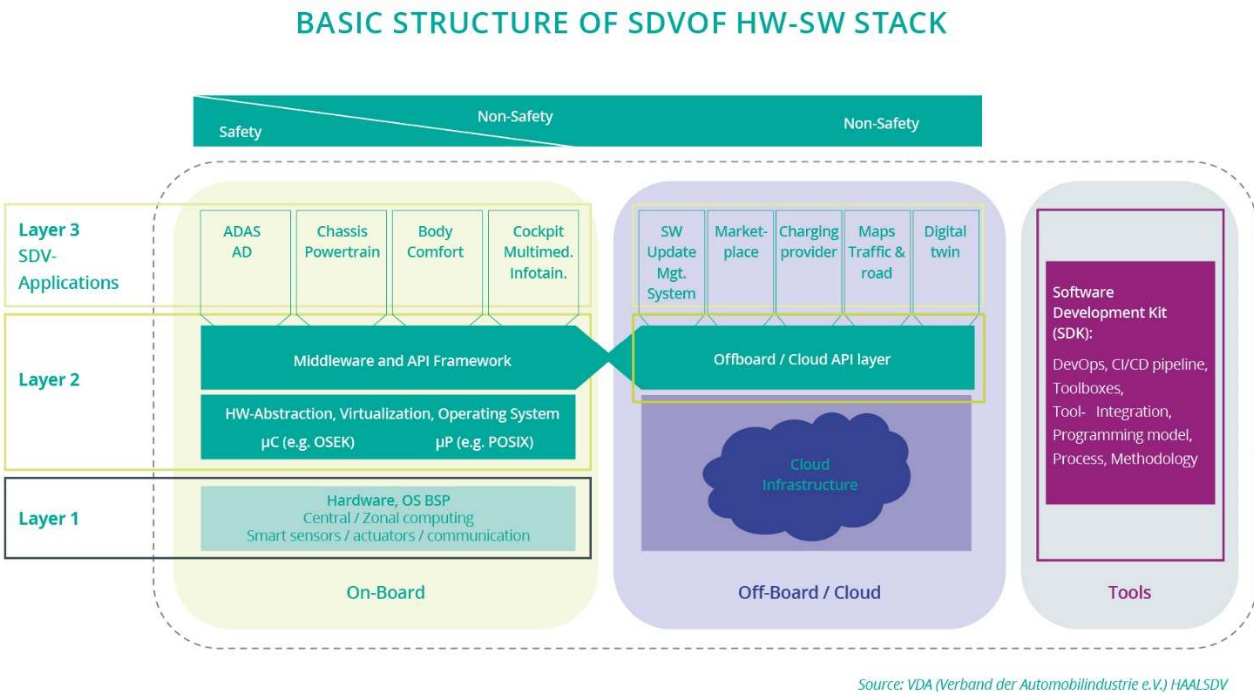


Figure 3.1-6: Basic structure of SDVoF HW/SW stack

The SDV structure abstracts the HW resources as processor resources, sensors, actuators, data coming from various communication media as well as data in the cloud (e.g. about traffic status) and the services build upon them for the application software used by vehicles, its drivers and passengers. This shall increase development efficiency and agility for user-relevant software applications. The SDVoF HW/SW-stack consists of the following layers:

- Layer 1 (SDV-Hardware): This layer includes all hardware components as automotive High-Performance Compute-platforms (HPC) and domain controllers and the SW abstraction of resources of the HPCs and domain controllers required for the largely AI-based software, serving several automotive domains with applications satisfying requirements like non-safety critical, safety critical, security, energy efficiency, etc.
- Layer 2 (SDV-middleware & hardware abstraction & OS): This software Layer 2 is the major focus of the SDVoF initiative. It consists of (mainly open-source and mostly non-differentiating as well as some closed source and/or differentiating) software building blocks, which connect the hardware layer with the applications layer to allow separated hardware and software development cycles necessary for software-defined vehicles of the future. This layer also ensures the safe execution of the applications of layer 3. It shall consist of software building blocks, which can be used also in existing OEM specific SW stacks. As many applications have on-board and off-board (cloud) parts, a service-oriented interface layer shall also exist in the cloud.

- Layer 3 (SDV-Applications): This layer contains the differentiating parts with applications in automotive domains such as infotainment, automated driving, advanced driver assistance functions, chassis and powertrain control, cockpit user interfaces, e-charging, routing, body-and comfort functions, etc. Many of these applications have parts on-board the vehicle and other parts off-board in the cloud.

To ensure agility, non-differentiating modular software building blocks shall be developed, which allow to build SDV software platforms or enhance existing company-specific SDV software stacks. A consensual definition of interfaces for these non-differentiating building blocks is highly important. The integration into the company (OEM or Tier) specific software stacks may require adding thin layers of company-specific software ("glue logic") to cope with the differences in the existing as well as future company-specific architectures. As the focus of the building blocks is in the non-differentiating area, they will reside in mainly layer 2, which consists of essential functionalities as virtualization, car (meta) operating system, HW/SW abstraction, the on-board middleware, API framework and cloud middleware.

Many of these software building blocks will be open-source in order to allow a resource sharing in the development of non-differentiating parts of these essential SDV SW platforms. The essential agility and speed in the development of these highly complex SDV software systems shall be achieved by applying the "code-first" principles. The "Code-first" approach has several goals: firstly, iterative development to sequentially improve building blocks and integrate them as soon as possible into industrial SW stacks. Standardization effort will only start, when building blocks have proven their functionality, robustness, and quality in real-world usage. Secondly, test driven development is a very good development methodology of an agile "Code-first" approach. It basically means, that the test procedures are designed and already implemented as part of the requirement specification. This also helps to significantly improve the quality of requirements and then of the code. Additionally, it supports the automation of testing.

Key RDI topics

- Development of a scalable, cloud-capable, and modular target architecture that supports decoupling of hardware and software and features a strong middleware layer consisting of building blocks. The building blocks of the SW platform needs to support current and future E/E architectures (domain-oriented, zone-oriented etc.). Most of the building blocks are expected under an open-source license in one of the participating initiatives.
- Development of a neutral SDK (software development kit) which allows software application developer to develop and test independently of hardware and operating system.
- Abstraction of the complexity of underlying hardware, middleware, kernel, interfaces, and drivers into simple to use and to re-use, robust, safe & secure APIs.
- Supporting and ensuring safety and security requirements using automotive standards for significant parts of the SDV (refer also to the SRIA chapter 2.4 MJ 3).

A high-level view of a roadmap is explained in more detail in the document[13] created and updated by the CSA Federate and shown in the diagram below. It will be continuously aligned with relevant initiatives in the European commission as well as with industry initiatives mentioned above.

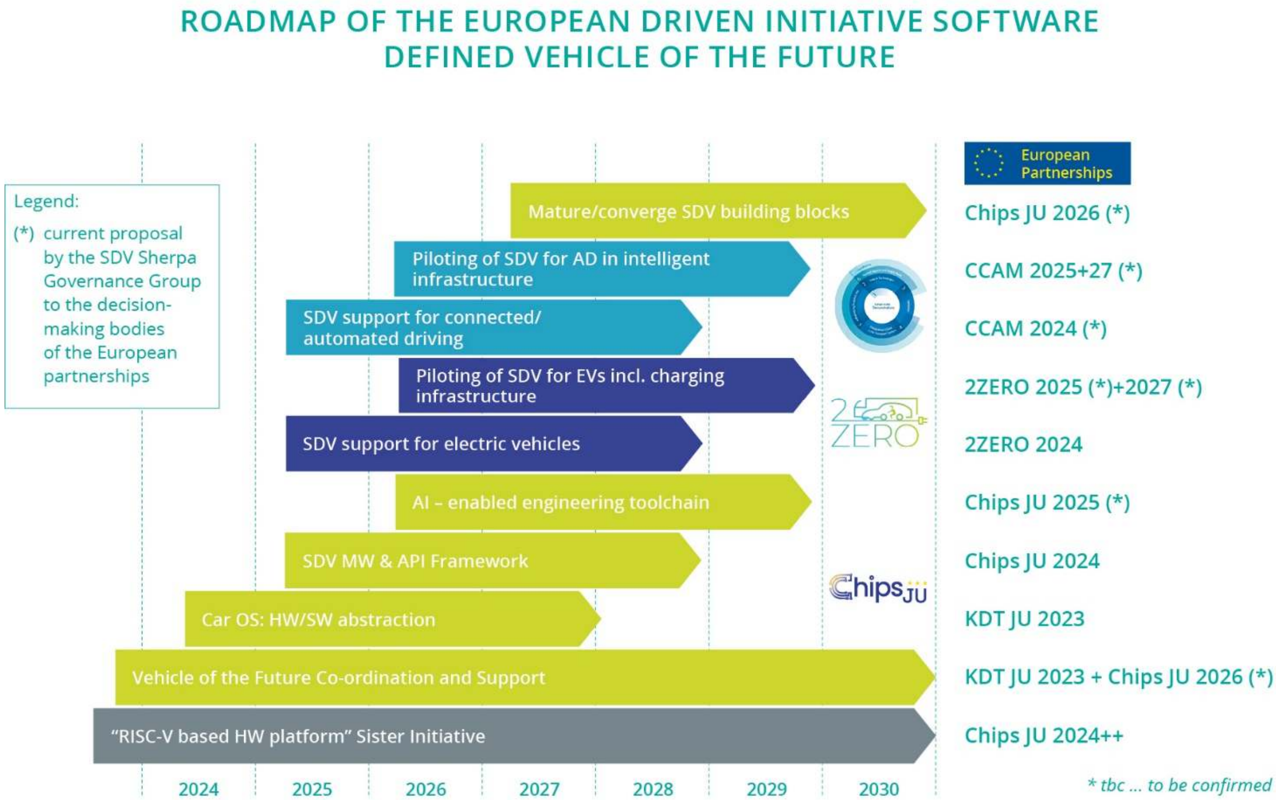


Figure 3.1-7 Roadmap of the European driven initiative Software defined vehicle of the future (SDVoF)

3.1.2.3 Major Challenge 3: Green deal: enable climate and energy optimised mobility

Worldwide efforts on the regulation of pollution and CO_2 emissions are leading to a strong increase in the electrification of vehicles, either with batteries ("battery electric vehicles", BEVs), vehicles with internal combustion engine and electrical engine ("hybrid electric vehicles", HEVs) or vehicles using fuel cells. Possible scenarios developed by BIPE[14] in France are shown in Figure 3.1-8. Depending on the evolution of regulations in particular, the split between the various energy sources could be significantly different between the scenarios in Figure 3.1-8. However, the most probable scenario is that of the "Green Constraint". But currently there are discussions about bans of fossil fuel engines even as soon as **2035,** therefore the data are expected to change.

[13] https://federate-sdv.eu/wp-content/uploads/2024/04/2024-04-12-SDVoF-Vision-document-ver017-final.pdf
[14] https://lebipe.com/

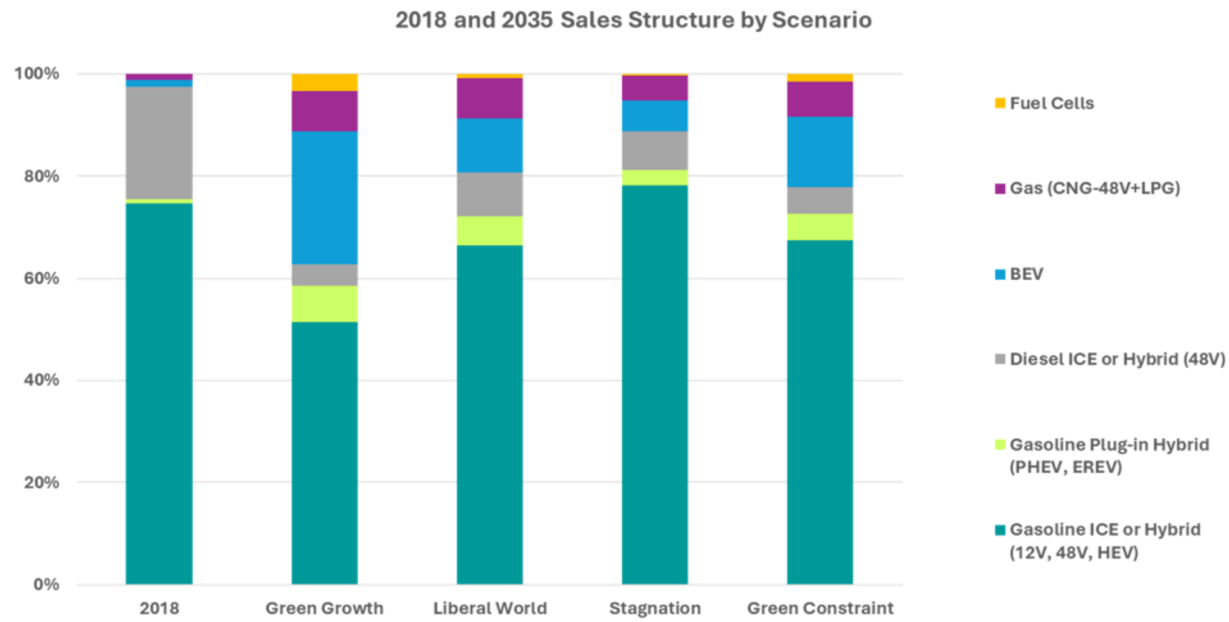


Figure 3.1-8 Sales structure of cars 2018 and 2035 (4 scenarios) (Source: PFA France 2018)

Looking in more detail at the difference between low voltage systems (particularly 48V) and high voltage systems, there are six important observations that can be made:

- All new cars in Europe will be electrified by 2035.
- The proportion of electrified cars in the world will reach about 70% in sales.
- Around one-third of the market will require on-board chargers for high voltage in the range 400–1000V.
- Fuel cell electric vehicles may play a significant role in long-haul trucks and trains, as well as in airplanes, ships, and drones by 2035.

Convergence of automotive and energy eco-system

Electrified cars are becoming an integral part of the European energy eco-system. Vehicle batteries will be used as intermediate storage for electrical energy (bidirectional charging). Vehicles will even produce electrical energy (photovoltaic panels on car roofs or from heat-pumps), which will be partially send to the grid. This has severe consequences for the requirements of the electrical and software systems in the vehicles. Their operation hours over the intended lifetime of the vehicle are significantly increasing. Up to now, the lifetime for the components of vehicles is about 8000 h, which reflects the hours a typical vehicle is driving. As the vehicle will soon become an integral part of the grid, its total number of operating hours will be the total number of hours per year (both running and connected-to-the-grid time) multiplied by the expected lifetime of the vehicle in years. This means an increase of lifetime of the effected components (as in the bi-directional charging or the necessary embedded control software) by a factor of 5 to 15. This convergence of the automotive and energy eco-system creates also large challenges in the validation. It requires significant use of digital twins in the simulation.

The expectation is that the overall electrification scenario will lead to massive changes in the supply chains and the distribution of competences. An important aspect in the CO_2 reduction is the speed of reduction. Selling new cars with best performance characteristics will only show an effect in the medium and long term. Any solution that improves the existing parc of 250 million cars in the world would have a much higher impact.

Car sharing, more efficient tires, teleworking and better lubrication can also add-up to significant and instantaneous CO_2 reductions. In this respect, the electronics industry should regard solutions that support rapid deployment of CO_2 reduction initiatives.

Key RDI topics

- Modular, flexible, and scalable platforms and electrical/electronic (E/E) architectures.
- Hardware upgradability and packaging that supports repair
- Exchange of existing components and systems by higher efficient modules.
- Reconfigurable and adaptable software architectures.
- Software updateability (including over-the-air, OTA).
- On-board technologies (devices, actuators and sensors, virtual sensors).
- Embedded intelligence (AI-powered and AI-enabled intelligence):
    - Control software, real-time capable algorithms.
    - Fault-tolerance, fail-operational concepts.
    - Cognitive vision.
- New control software is required to take full advantage of new solid-state batteries so that they have an extended lifetime, as well as a higher driving range in vehicles. Other trends are SIBs and high silicon content. As lifetime is key for batteries used in mobility systems, tools for accelerated lifetime testing, diagnostic systems as well as control systems that can extend the lifetime and limit degradation, are essential for the success of electrified green mobility. New power electronics based on silicon carbide (SiC) and gallium nitride (GaN) devices are needed to ensure energy-efficient operation. AI and model-predictive control algorithms, supported by high-performance, multi-core, real-time operating systems, has to offer the necessary intelligence based on ultra-low power/high-performance control units.
- Similarly, advanced control methods for fuel cell-based vehicles (mainly in trucks and buses) that both minimise degradation and maximise efficiency are crucial. For example, predictive control schemes that take into account forecasts on e.g. route, traffic, weather, etc. are necessary. State-of-health monitoring systems (virtual sensors) as well as adequate new sensors to measure the operating conditions within fuel cells without negatively influencing their operation, are required.
- For both electric battery and fuel cell-based mobility, new safety concepts using (AI-based) IoT diagnostics must ensure the safety of these systems, especially in accident situations.
- Low environment impact of new technologies in terms of energy consumption at production, deployment, use and end of life treatment.
- Reduction of sensible materials such as rare earth materials.

- Recyclability of electronics: the environmental impact needs to be considered (eco-design, circularity…) because it requires an ecosystem approach, which includes the mobility application industry
- Efficient and fast charging and filling of alternative energy into green vehicles is another critical research topic. Being the glue between Transport & Energy Sector, "Plug and charge" is an essential topic.
- The conversion of renewable energy into green energy, as electricity stored in vehicles or H_2 or alternative fuels, also needs efficient electronics with real-time embedded software with energy management SW communicating to the power grid to minimise the need of new charging/filling infrastructure, which is one of the cost drivers limiting the speedy success of green mobility.
- Smart Battery: with the batteries in electric vehicles being the most expensive and life-time critical parts, future battery systems will be equipped with more sensing technology, intelligence, and communication systems to monitor their own health and record their lifetime dataset. This enables better usage of the batteries as well as optimised second life concepts.
- Power electronics (fast-switching elements, wide bandgap materials, low power, etc.).
- Predictive diagnosis and maintenance (including recovery strategies, fault detection and localisation, surveillance sensors, etc.).
- Cloud/edge/fog processing approaches.
- Standards, including communication and interoperability standards, electromagnetic spectrum, and bandwidth management, charging units, car access systems, etc.
- Proof of robustness and trustworthiness of architectures and quantification of the operational risks.
- Collaborative and self-organised multi-agent systems, e.g. in logistics applications also covering cooperation between land and air vehicles.

By 2050, 67% of the population is expected to live in urban areas. As cities become bigger and smarter, this trend will lead to new opportunities for tailored and specialised vehicle design specific to urban users, including the needs and operations of commuters, as well as ride-hailing and last-mile delivery. New vehicle concepts and ECS-enabled architectures will lead to flexibility, scalability, and modularity - while featuring safety, security, and reliability - to ensure urban-readiness (appropriate range, compatibility with charging infrastructures, ease of parking and operations, etc) in all kinds of urban and suburban areas, most likely with different implementation levels of infrastructure and smart technologies. Additionally, it is assumed that these vehicles will not have to be designed for high-speed operation and long range and can easily be charged sufficiently fast and comfortably to meet the daily needs of urban and suburban mobility usage scenarios. This aspect may also include sharing concepts, and consideration should also be given to use by the elderly and disabled.

Another important aspect is the need for reliable and efficient wireless communication technology combined with different types of sensing systems to achieve efficient traffic and increase safety as well as reduce fatalities. This further sets requirements on components and systems for wireless communication to achieve ultra-high reliability and resilience as well as to meet challenging performance and latency demands.

The challenge targets the following vehicle categories:

- Passenger cars (including light four- wheelers, M1/M2 category).
- Trucks (including power-driven vehicles having at least four wheels used for the carriage of goods (N1), agricultural and forestry tractors, and non-road mobile machinery (T)).
- Ships.
- Airplanes.
- Motor vehicles with less than four wheels (L category).
- Off-road vehicles (G).
- All kind of unmanned air vehicles (such as drones).
- All kind of manned light air vehicles.
- Special-purpose light vehicles (air, land, water).

Creation of synergies between different transport systems can bring cost advantages (automotive – air – freight)

The developments described above ask for innovative mobility solutions in the years to come, affecting European society:

- Urban light personal and small freight mobility (including innovative micro-vehicle designs suitable for urban/suburban commuters' needs, with an option for usage within shared mobility schemes. Such micro-vehicles would also be capable of interfacing with urban collective transport systems (i.e. easy access to buses, trams, and trains for last-mile transfers to achieve full intermodality).
- Light and flexible multi-passenger vehicles (e.g. collective, or individual, owned or shared up to M1 category) with robust safety measures for passengers and vulnerable road users, and including specific features to facilitate shared use such as autonomous-capable vehicles with automated relocation to charging points or areas with insufficient vehicle density.
- Long-haul and right-sized vehicles and tailored ECS for commercial uses, such as for long distance, last-/first-mile delivery, construction and maintenance support, which are suitable for urban scenarios.
- Environment friendly mobile machinery to optimise harvesting (and reduce accidents).

3.1.2.4 Major Challenge 4: Digitalisation: affordable and safe automated and connected mobility for passengers and freight

Regain European sovereignty in embedded software for mobility

All of the four mobility megatrends (electrification, autonomy, connectivity, shared mobility) highly rely on leading-edge software. Especially the car industry needs to face this new software-driven value chain. A typical modern vehicle likely has a software architecture composed of five or more domains (body, chassis, ADAS/AD, powertrain, infotainment, …), together comprising hundreds of functional components in the car and in the cloud. Currently, OEMs are constructing a complex software architecture with various software providers ending up in a complex scenario with a broad set of development languages and interfaces/APIs, operating systems, and software structures/elements. Up to now no single software platform (middleware) on the market can meet the requirements of Connected, Cooperative and Automated Mobility (CCAM). The domain-based hardware architecture is rapidly evolving towards a zonal architecture, where the software architecture, implementation and the corresponding development environment needs to be compatible. A middleware layer has to separate the hardware layer from the software layer and allow faster time-to-market to standardised interfaces.

The work on the mobility challenges will help to protect the sovereignty in the digital heart of mobility devices (passenger cars, trucks, …). In the past, Europe had a leadership in the embedded software in all kinds of vehicles. This is endangered by the push of the big global IT companies into the embedded IT systems of the vehicles. Europe has to protect its leadership in creating a next generation of middleware and operating systems prepared for the significantly increased complexity of CO_2-neutral automated and assisted vehicles.

Europe envisions fatality-free transportation as well as seamless mobility choices (including multi-modal services) for its citizens, particularly in view of the aging society. Moreover, the transport industry in Europe in general aims to maintain its leading position by offering sustainable solutions for safe and green mobility across all transportation domains – automotive, avionics, aerospace, maritime (over water as well as under water transport) and rail. Europe's strength lies in its established expertise in developing complex electronic components, cyber-physical systems, and embedded intelligence. However, hurdles related to autonomy, complexity, safety, availability, controllability, economy, and comfort need tackling as automation increases. The overall vision is to achieve always-connected, safe and secure, cooperative and automated transportation systems (passenger transport, cargo and freight movement) using highly reliable and affordable European-made electronic components and systems. These systems will also incorporate new human-machine interaction technologies.

For the road sector, objectives are (a) to reduce the number of road fatalities and accidents caused by human error to zero by 2050, (b) to ensure automated transport introduces no additional road fatalities, and (c) to decrease validation costs down to 50% of development costs from the current 70–80%. Key Digital Technologies are crucial for achieving these goals and supporting the ambitions of the Horizon Europe Partnership on Connected, Cooperative and Automated Mobility (CCAM). Collaborations in and across industrial domains, learning from operational field data, and joint strategic actions are essential. No single entity can singlehandedly manage these significant R&D efforts.

For waterborne transport, as automation progresses ships will become fully connected, globally. Remote vessel monitoring is already possible, allowing for condition-based maintenance. Building on increasing onboard automation, the remote operation of vessels will become possible, eventually moving towards full autonomy for vessels. The wider use of unmanned autonomous vessels (UAVs) – either aerial, underwater or on the surface – will increase the flexibility and energy efficiency of operations. For avionics and aeronautics, automation is paving the way for more efficient flight management, reducing pilot workload, and enhancing predictive maintenance, which collectively promises safer and more efficient air travel. Additionally, Urban Air Mobility (UAM) is revolutionizing cityscapes, with automation enabling flying taxis and drone deliveries, presenting a transformative solution to urban congestion, and offering a new dimension to urban transport. For the rail sector, automation is driving the development of driverless trains, optimizing track usage, enhancing scheduling accuracy, and ensuring safer and more punctual journeys for passengers.

Connected, cooperative and ultimately automated mobility will play a central role in shaping our future lives. ECS will enable different levels of partial, conditional, highly and fully automated transportation, posing new challenges for traffic safety and security in mixed scenarios where vehicles with different automation levels coexist with non-automated vehicles. Reliable and fail-operational approaches for tele-assistance and tele-operation need to be considered as well to remove the safety drivers from the cabin. Both stepwise automation ("conversion design") and full automation (Society of Automotive Engineers, SAE, level 5: "purpose design", e.g. a people mover in a structured environment), have been developed to the prototype level and are being deployed in field-operation tests right now. Nonetheless, these solutions need to be further explored and improved towards safe and reliable capabilities in extended operational design domains. Additionally, cross-fertilisation with other industrial domains such as Industry 4.0 and Industry 5.0[15] and evolving communication technologies (5G/6G) is essential.

As the proportion of electronics and software, considered as a percentage of the total construction cost of a vehicle, increases, so does the demand for the safe, secure, reliable, and unhackable operation of these systems. In addition, privacy protection is paramount for car owners, drivers/operators and passengers as well. This demands robust and fail-operational technologies that deliver intrinsically safe operation and dependable fall-back position from component to subsystem and provides a solution for problems in interaction with the cloud. Also, multi-core-based platforms and sensing devices, combining advanced sensing in harsh conditions, novel micro- and nano-electronics sensors, advanced sensor fusion and innovative in-vehicle network technology are needs. While CCAM promises significant benefits, it also faces numerous challenges to be faced. These challenges are manifold since the include technical, safety & security, legal but also social challenges.

Safety and security for humans are essential. ECS systems as health monitoring, drowsiness monitoring, etc. will focus to increase safety and healthiness of vehicles passengers. Security of the passengers is another important RDI area, as increasing communication with various cloud systems, complex infotainment systems and powerful AD sensors open many intrusion doors.

[15] https://research-and-innovation.ec.europa.eu/research-area/industrial-research-and-innovation/industry-50_en

Key RDI topics:

Technology challenges

- **Complexity of algorithms**: Developing robust, reliable algorithms for autonomous driving that can handle a wide range of scenarios and conditions is highly complex. This includes perception, decision-making under uncertainty, and control algorithms in multi-agent and multi-modal settings.
- **Sensor limitations**: Sensors like LiDAR, radar, and cameras can be limited by weather conditions, lighting, and other environmental factors. Ensuring reliable sensor performance in all conditions is the key to succeed.
- **Data processing and integration**: Autonomous vehicles generate vast amounts of data that need to be processed in real-time. Integrating data from multiple sensors and sources while ensuring low latency is critical.
- **Infrastructure Requirements**: CCAM requires significant investment in infrastructure, such as high-definition maps, roadside units (RSUs), and vehicle-to-everything (V2X) communication systems. Implementing and maintaining this infrastructure is costly and complex; prioritization is required along with robust connectivity services.
- **Artificial intelligence (AI)**: Reduction of the computational complexity of AI algorithms, advancements in energy-aware hardware design, including low-power processors, energy-efficient sensors, and power management techniques, along with edge-cloud collaboration, resource management, and continual learning / adaptability strategies are highly required.

Safety and Security Challenges (refer also to SRIA chapter 2.4 Major challenge 3)

- **Cybersecurity: CCAMs are vulnerable to cyber-attacks, which can have severe safety implications.** Ensuring robust cybersecurity measures is essential.
- **Safety assurance**: Proving the safety of autonomous systems to the level required for public deployment is difficult. This includes extensive testing, validation, and establishing safety standards.
- **Fail-safe mechanisms**: Developing fail-safe mechanisms and redundancy in autonomous systems to handle system failures or unexpected situations is a significant challenge.

Regulatory and legal challenges (refer also to Chapter 2.4, major challenge 4)

- **Regulatory frameworks and liability**: Creating consistent regulatory frameworks that can keep pace with rapid technological advancements as well as determining liability and allowing for interoperable and fast deployment across Europe is challenging.

Ethical and social challenges

- **Ethical decision-making**: Autonomous vehicles must be programmed to make ethical decisions in situations where harm is unavoidable. Creating consensus on these ethical frameworks is difficult.
- **Public acceptance**: Gaining public trust and acceptance of autonomous vehicles is crucial. This involves addressing fears and misconceptions about the safety and reliability of the technology.
- **Impact on employment**: The widespread adoption of autonomous vehicles may impact jobs, particularly in sectors like transportation, logistics and digital industry.

Environmental and Economic Challenges

- **Energy consumption**: Autonomous vehicles, especially when equipped with multiple sensors and communication systems, can have high energy demands. Balancing energy efficiency with performance is necessary.
- **Economic viability**: The high cost of developing, testing, and deploying autonomous vehicles and the necessary infrastructure can be a barrier to widespread adoption. Ensuring economic viability for manufacturers, consumers, and municipalities is challenging.

Interoperability and standardization

- **Interoperability**: Ensuring that autonomous vehicles from different manufacturers can communicate and operate seamlessly with each other and with infrastructure is critical. This requires standardization of communication and data exchange protocols and systems.
- **Standards development**: Developing and implementing standards for autonomous vehicle performance, safety, and communication is a complex and ongoing process that requires coordination among industry stakeholders, regulators, and standards organizations.

Addressing these challenges requires collaboration among technology developers, policymakers, regulatory bodies, and the public to ensure that CCAM can be safely and effectively integrated into *society.*

Key elements of ECS for cars that need to be developed are shown in Figure 3.1-9.

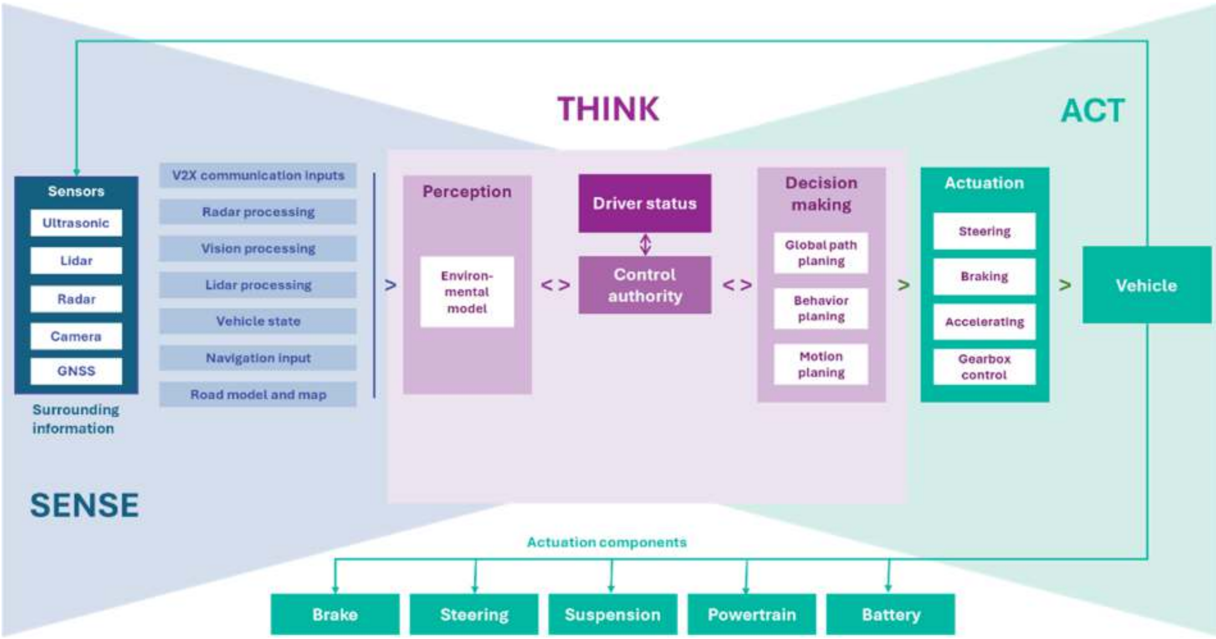


Figure 3.1-9 ECS-based components of automated vehicle

Key focus areas

Based on the above identified challenges, the following research priorities and development & innovations areas have been identified to boost the maturity and deployment of CCAM:

Robust Perception and situational awareness: Reliable perception systems are essential for enabling robots to interpret and navigate complex environments effectively. Dependable and affordable environment perception and localisation sensors, and V2X communication. Attention should be paid to sensor interference, more in particular the robustness of sensors to environmental conditions, to interference by other sensors and to malicious interference.
Integrated sensing and communication systems are necessary to further evolve towards fully automated transport and safeguarding VRU's (Vulnerable Road Users) in all type of traffic and weather situations (especially also at night or in adverse weather).

Real-time collaborative decision-making under uncertainty is essential for enabling autonomous vehicles to respond quickly to dynamic environments and perform tasks with precision. Research efforts should focus on optimizing non-AI and AI algorithms for low-latency execution on edge devices, leveraging techniques such as model parallelism and hardware acceleration. Edge computing architectures must be designed to minimize processing delays between data acquisition and actuation.
Hardware-software Co-design, digital twinning and continual learning: The synergy between hardware and software components is critical for maximizing the performance and efficiency of CCAM systems. Research should explore co-design methodologies that tailor hardware architectures to the specific computational requirements of AI algorithms used in robotic applications. The ability of robots to adapt and learn from their interactions with the environment is crucial for achieving long-term autonomy and

versatility. Advancing from ego-vehicle optimization to multi-agent settings optimization including edge-cloud communication is a logical step forward.

Real-time processing: Centralised service/function-oriented hardware/software architectures, including open APIs, for road vehicles, ships, trains that are supported by the cloud and edge computing via 5/6G along with dependable and reconfigurable hardware and software, including remote access and Over-The-Air (OTA) software upgrades. Hardware and software platforms for control and higher performance in-vehicle networking units for automated mobility and transportation (including support for AI), for example the usage and adaptation of IoT integration platforms, also for automated and connected environmentally friendly vehicles. New developments towards higher performance and efficiency are crucial: These are also required to ensure the reliability and safety of the power electronic components and systems for the drivetrain and charging systems, as well as for steering, breaking/suspension/air condition control in automobiles, trains, ships, and flying equipment.

Advanced AI-based and predictive control methods to create new active safety paradigms for the next generation of automated vehicles, e.g. to operate beyond the boundaries typical for the current generation of stability controllers and chassis control systems, when this is needed to prevent road accidents. This also implies a redefinition of the software architecture to integrate the chassis control and driving automation functions. The current generation of chassis control systems – which will be used in the first generation of automated vehicles – is designed to keep the vehicles within limits that are desirable for human drivers, but further research is required to assess the new active safety options (e.g. race driving techniques to prevent crashes in emergency conditions) enabled by driving automation. Multi-functional cross-domain and preview-based control software implementations, e.g. powertrain and brake-by-wire controllers concurrently managing energy efficiency, active safety, and comfort aspects (e.g. powertrain torque distribution to induce desirable pitch in braking, compensation of the longitudinal acceleration oscillations caused by road irregularities through road preview-based control of the powertrain torque).

(Predictive) health monitoring and lifetime analysis for the perception and control systems (including all required sensors, V2X systems and localisation systems) and AI components of (highly) automated vehicles used in the operational phase.

These research priorities result in completely new hardware and software architectures for the control systems (non-AI, AI-based) due to exploding sensor data volumes, dynamic adaptability, multi-agent requirements as well as new power saving hardware and software components. This requires a decoupling from hardware with drivers and operating system, from the application components, adaptive AI-based routing and control algorithms, multimedia components and many more, and at the same time a co-designed process of interface developments. These demands directly impact the Chips JU strategy on software-defined vehicles (SDV)[16]. As CCAM methods and technology advances should be application independent per se (of course, capable of being tailored to specific applications), collaborations with the AI and robotics community should be strengthened and implemented, e.g. with the AI Data Robotics Association (ADRA)[17] or the US National Robotics initiative[18].

[16] https://federate-sdv.eu
[17] https://adr-association.eu
[18] https://hichristensen.com/post/roadmap-2024/

3.1.2.5 Major challenge 5: Edge2cloud mobility applications: added end-user value in mobility

The rapid advancement and integration of Cloud-Edge-IoT (CEI) technologies are primarily driven by an increasing demand for digital twins, immersive communication, and secure mobile robotics, alongside the need for low-latency and real-time capabilities at the edge of networks for a multitude of innovation fields. These technologies are crucial in addressing the challenges posed by energy shortages and climate change, as they promote energy and resource efficiency through new hardware and AI-based optimizations. At the same time, they provide advanced intelligence and data-based decision making capabilities deep in the networks of technical systems of systems. These leads to powerful mobility services distributed between cloud and vehicle SW. Off-board SW technologies differ from classical vehicle SW technology and need to be integrated into mobility application architectures. Cloud-native architectures are moving partially also into vehicle SW in these new mobility applications.

As such, these innovations not only meet the technical demands of a modern digital infrastructure but also offer competitive advantages and open new business opportunities within a globally competitive market. Transport and mobility is a major application sector of CEI technologies since it aims at vision of zero emissions, zero accidents and zero hurdles between modes or a sustainable, safe and convenient personal mobility experience or freight service.[19] The growing use of CEI systems in transport and mobility presents a major opportunity for European companies to increase their cloud capacities and improve their edge computing technologies. This could help them compete with international rivals over time. However, for now, the dominance of international hyper-scalers in providing cloud services may impact the influence and control of established European players.

Some general potentials and issues of applying CEI in the transport and mobility sector need specific attention:

Opportunities:

- Data critically determining safety, privacy, and efficiency of transport operations can be effectively processed with low latency and high security on edge platforms close to sensors and actuators.
- Operations of strategic relevance for transportation, such as planning, modelling, data analysis, software engineering, maintenance, AI computations, and the management of software updates, including the potential resale of data, can be effectively centralized in the cloud.
- CEI technologies pave the way for innovative services in the transport sector, such as shared transportation platforms and automated vehicles, enhancing system efficiency and user experience.

Challenges:

- Concerns about data sovereignty and privacy may hinder the deployment of CEI technologies, particularly in the public transport sector, where the mass transfer of data to international servers with limited oversight is faced with scepticism from users.
- The ongoing debate over data ownership, particularly in the public sector, about whether data should be generated and managed by public authorities or outsourced to private service providers applies to transport operators emphasizing the need for data sovereignty in governmental data management.
- The transport sector faces a risk of vendor lock-in, with limited alternatives for European IoT companies and users in selecting cloud providers, which could stifle innovation and competition.

---

[19] Deliverable 3.1 - CEI ecosystems overview with the value chain adopter groups. EU project UNLOCK-CEI, 2022.

- Cybersecurity is increasingly challenging due to the growing communication channels from the vehicle to the outside world.

As the functional scope of transportation sector is quite divers considering transport modes in passenger and freight transport, the field of IT-solutions and thus conceivable CEI applications is very large. The CEI applications may concern transport networks, vehicles, traffic management, mobility services, data platforms, logistics applications, fleet management, planning applications etc. An efficient Cloud-Edge-IoT (CEI) infrastructure for transport and mobility applications requires multiple innovations across several functional layers, each with specific needs:[20]

- **Design Layer**: This layer must accommodate real-time processing of large data volumes, such as those from traffic management systems. It requires substantial computational power to handle data from diverse road users including cyclists and pedestrians, not just vehicles. Integrating data protection, privacy, and adherence to regulatory and governmental standards is crucial to ensure the system's legitimacy and trustworthiness.
- **Installation Layer**: As the number of sensors in traffic management increases, the installation layer must address the challenge of power supply, possibly smart charging utilizing e.g. lamp poles, interacting with the power grid and exploring new energy sources. Efficient connections between new sensors and edge computers need to be established, and the financial implications of installation must be discussed, ensuring robust communication capabilities between different city departments.
- **Operation Layer**: The operational functionality must manage dynamic traffic volumes effectively e.g. through traffic lights and controls) or provide efficient and seamless transfer between transportation means (e.g. though integrated ticketing), maintaining high data quality and security even in adverse conditions through the use of AI and appropriate sensors. Data should be real-time, accurate, and secure, yet openly accessible where appropriate. Proper sensor placement and the ability to retrieve data from multiple computers are essential for operational integrity.
- **Value-added Supplements**: Innovative features such as two-way communication between automated vehicles and traffic signals, and enhanced capabilities for rapid incident or adverse weather detection, represent value-added layers that leverage CEI technology to improve traffic management and safety. The provision of content for entertainment systems is another way of adding value through CEI.
- **Maintenance Layer**: Traffic control equipment typically has a long lifecycle, often misaligned with the faster modernization cycles of other computing equipment. Maintenance practices should consider this discrepancy to ensure consistent performance and easy upkeep.
- **Disposal/Upgrade Layer**: Given the long lifespan of traffic control systems, a strategic approach is necessary for integrating and replacing legacy systems. Complete system replacements are rare; therefore, a planned, phased integration or upgrade strategy is crucial to maintain continuity and efficiency in traffic management.

To effectively implement Cloud-Edge-IoT (CEI) technology in transport and mobility applications, specific technological advancements in ECS necessary namely:

- **Enhanced Computing Power at the Edge**: To support real-time processing of traffic data, edge devices installed close to traffic sensors and actuators must possess significant computational capabilities.

[20] Deliverable 3.2 – Sector-specific service requirements, data flows and revenue streams in Cloud-Edge-IoT value networks. EU project UNLOCK-CEI, 2022.

This allows for immediate data processing and response, crucial for e.g. connected and automated driving systems.

- **Advanced Connectivity and Networking:** Reliable, high-bandwidth and low-latency communication networks are essential to manage the continuous data exchange between vehicles, traffic management systems, and infrastructure. This includes vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications, which are foundational for automated driving technologies.
- **Smart Data Management Systems**: Implementing comprehensive sensor systems requires robust data management to ensure privacy, security, and accuracy. Systems must be designed to handle large volumes of data generated in real-time from multiple sources across the transport network.
- **Scalable and Adaptive Infrastructure:** Infrastructure must be capable of scaling effectively to accommodate an increasing number of automated vehicles and complex traffic management tasks. Modular infrastructure design allows for expansions and upgrades as traffic technology evolves.
- **Intelligent Sensor Networks**: Deploying smart sensors throughout the transport infrastructure that can reliably function under various environmental conditions is critical. These sensors gather critical data for traffic management and vehicle operation, feeding into edge computing systems for rapid processing.
- **Integration of AI and Machine Learning:** AI and machine learning are pivotal in enhancing traffic management services, such as those based on digital twins of road infrastructure. These technologies enable predictive analytics, traffic flow optimization, and incident detection, all processed locally to reduce latency.
- **Interoperability Standards**: Standardization across all devices and systems ensures seamless communication and functionality within the interconnected transport environment. This supports the integration of various manufacturers' vehicles and infrastructure components.
- **Robust Cybersecurity Measures**: Transport systems involve critical safety and operational data, necessitating stringent security measures to protect against cyber threats. This includes encrypted communications, secure data handling protocols, and regular security audits.
- **Energy-Efficient Design:** Given the extensive use of sensors and computing devices in transport applications, energy efficiency becomes crucial. Utilizing renewable energy sources and developing low-power devices can significantly reduce the environmental impact.
- **Maintenance and Upgrade Flexibilit**y: Transport and mobility systems require maintenance protocols that minimize downtime and disruption. Additionally, the system should be designed for easy upgrades to incorporate new technologies and standards without complete overhauls.

Future research and innovation activities on Cloud-Edge-IoT (CEI) technologies in transport could prioritize large-scale pilot projects, particularly integration AI to enhance real-time traffic management and vehicle automation. Key areas include:

- **AI-driven Traffic Management:** Developing systems that utilize real-time data from individual road users to dynamically control traffic situations. This involves AI algorithms capable of processing large volumes of data at high speeds to improve traffic flow and safety.

- **AI-supported automated tolling systems**: Developing systems for vehicles which automatically pay tolls across all roads in Europe and next to Europe without driver interaction. It requires significant work and alignment in the cloud. Here the automotive industry as well as the road operators across Europe have to cooperate. Regulatory support may speed up this difficult alignment process and the solutions for the development challenges
- **Automated charging systems**: Developing systems automated charging and billing. Next steps may include systems, which solve the problem, that low speed charging often results in high parking costs after the completion of the charging for customers, who have no opportunity to charge at home. Systems, which automatically move cars from charging positions to parking positions may offer a solution, alternatively automatic repositioning from charging infrastructure from one parking position to a next parking position after completion of charging could solve the problem. Here the automotive industry as well as the charging provider industry across Europe have to cooperate. Regulatory support may speed up this difficult alignment process and the solutions for the development challenges. Without such systems, the acceptance of electrical vehicles for the large part of the European population without a fixed parking space will be very slow, which may significantly endanger the transition to CO_2 free mobility.
- **Automated Parking-fee payment systems**: Developing systems for vehicles which automatically pay parking fees across all public parking garages in Europe and next to Europe without driver interaction. It requires significant work and alignment in the infrastructure and the cloud. Here the automotive industry as well as the parking garage operators across Europe have to cooperate. Regulatory support may speed up this difficult alignment process and the solutions for the development challenges
- **Road weather condition info systems**: Automated vehicles would benefit from information about wet roads, icy roads, slippery roads. This information can be derived from a wheel speed analysis from other cars, whose results are collected from road operators. They can merge the results with weather data, dedicated infrastructure systems data and weather forecasts as well as AI based predictions. Again, a collaboration between automotive industry and road operators is necessary
- **Bidirectional charging to stabilize the electrical grid**: Developing systems for vehicles which can utilize the vehicle batteries to stabilize the grid. It requires significant work and alignment in the cloud. Here the automotive industry as well as electrical energy providing industry across Europe have to cooperate. Regulatory support is essential.
- **Local Level Traffic and Pedestrian Management**: Projects that focus on detecting and controlling traffic and pedestrian movements at the local level to facilitate real-time responses and enhance connected and automated driving systems.
- Many more similar SDV application, which require in-vehicle development combined with development in other industries are possible. Many of these applications will require an App store for vehicle SW purchase-able via OTA.

All these applications are candidates for applications in SDV, which offer a significant increase in comfort for customers as well as an important societal value. They have in common, that different industries have to work aligned together with the automotive industry. The automotive industry cannot solve these problems alone. In most of the applications regulatory support may be required too for speeding up the process. Associations as CCAM and 2Zero are essential in the coordination of the different industry segments.

Common technological building blocks for these applications are:

- **Automotive Edge Platforms**: Creating advanced platforms for software-defined vehicles that can process data on the edge. This would support the higher computational needs of automated driving technologies and enable more responsive vehicle behaviour.
- **Strategic and Tactical Decision Support:** Projects should cover the development of decision-support tools that use global optimization in the cloud for strategic planning and edge computing for tactical responses, such as managing crossroads during incidents or heavy traffic.

Obviously so, CEI is closely to the other Major Challenge for ECS in transport and mobility.

3.1.2.6 Major Challenge 6: AI enabled engineering tool chain: agile collaborative SDV SW development and SDV as well as ADAS/AD validation

In the automotive industry, the significance of both hardware and software continues to grow as vehicles shift toward autonomy, electrification, connectivity, and service-oriented models. To guide software-defined vehicle development, several key principles emerge:

- **Open Source**: Embrace open-source practices, fostering collaboration and transparency. Open-source software allows for community-driven innovation and accelerates development.
- **Code-First Approach**: Prioritize software development, recognizing that it drives value creation. A code-first mindset ensures agility and responsiveness to market demands.
- **Communities of practice**: Maintaining well-integrated bundles of HW/SW and making them widely available will strengthen the European innovation ecosystem, backed by start-ups and academic spin-offs.
- **Agility**: Adopt agile methodologies to iterate quickly, respond to changing requirements, and deliver features efficiently. Agile practices enhance flexibility and adaptability.
- **Keep It Simple**: Simplicity is paramount. Streamline software design, avoiding unnecessary complexity. Simple solutions are easier to maintain and troubleshoot.
- **Hardware/Software Abstraction**: Decouple hardware and software layers. Abstraction enables flexibility, allowing software to run seamlessly across different hardware platforms.
- **Service Orientation**: Design software with a service-oriented mindset. Prioritize modular components that can be independently updated, enhancing scalability and maintainability.
- Significant usage of **artificial intelligence**.

However, the European automotive industry faces intensified global competition due to non-EU manufacturers' early adoption of software-driven strategies. Large tech companies and hyper-scalers, leveraging substantial software budgets and indirect business models, are already dominating specific domains. Additionally, significant state aid in East Asia facilitates rapid market entry for new companies.

These challenges require new development and verification & validation processes and tools supporting the collaborative, agile open-source based development of the necessary SW stacks as described in major challenge 2.

The tool chains are mostly called CI/CT/CD pipelines and must solve the following challenges:

- Multi-party collaborative SDV development environment ("Just-in-Time (JiT) for SDV")
- Definition of the attributes for artefacts exchange between contributors and maintainers / distributors
- Provide the tools needed in an integration Sandbox for SDV SW stacks
- Test Strategies, tools and Sandbox for SDV SW stacks in the various SDV domains
- Prepare example guidelines for all IT aspects related to SDV
- Prepare example guidelines for all legal aspects related to SDV (e.g. for ADAS, AD, Data act, AI Act, SOTIF, ISO
- Tools and processes covering the full Product Life-Cycle Support of SDV

Multiparty collaborative SDV development environment

In the automotive industry typically, several organizations work together to develop and maintain complete SDV SW stacks. In the non-differentiating layers of the SW stacks many open-source building blocks are used. Classical processes using different CI/CT/CD pipelines and transfer code and executables via email or similar between the stakeholders. This leads to long time periods from SW changes until the SW is available in a tested in released form to be used by customers. These long time periods are unacceptable any more especially when solving safety critical bugs in ADAS/AD features found in the field. Therefore new cooperation models supported by tools are required to allow just-in-time SW integration and fully automated tests specifically adjusted to the modified modules in order to reduce the time from code-change at a Tier until the SW is ready for a vehicle from several month to maximum a week (Figure 3.1-10).

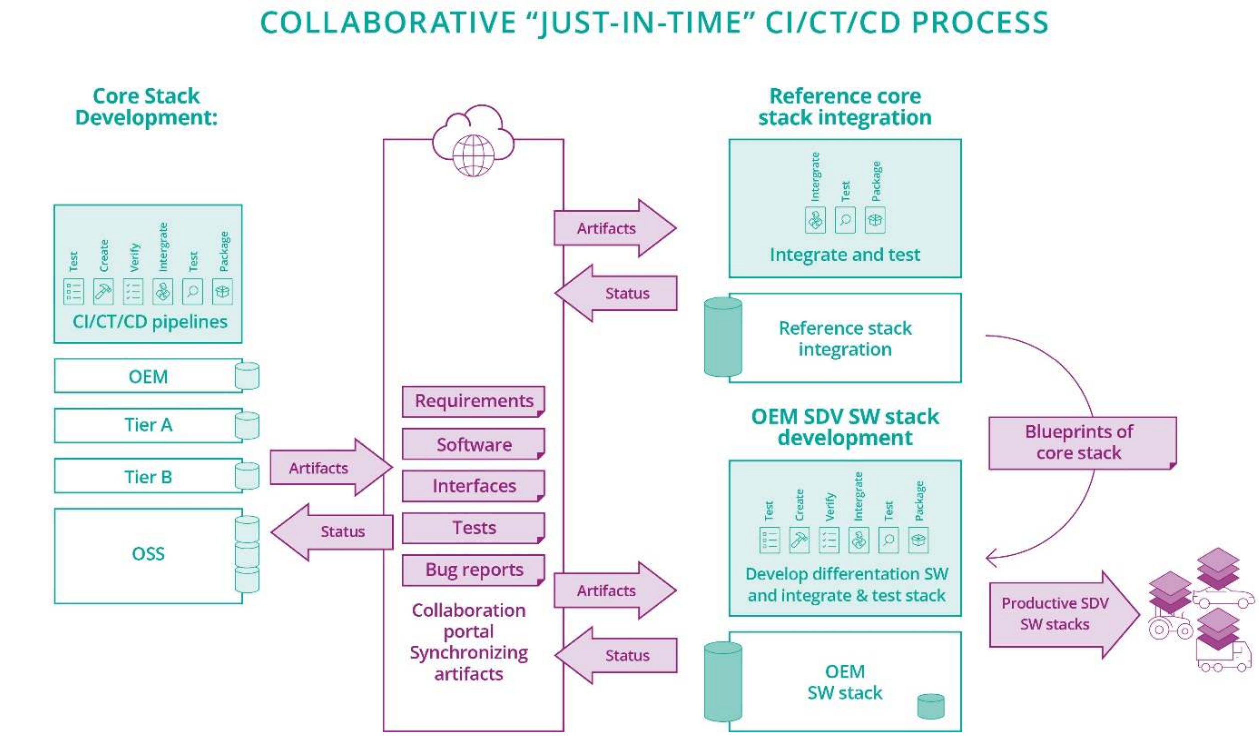


Figure 3.1-10 Collaborative « just-in-time" CI/CT/CD process (Source: Continental)

Key RDI topics

- Create and automatically install CI/CT/CD (- configuration as code -) build from open source tools (building blocks)
- Automated synchronization of building blocks with all necessary artifacts (including the test information from the out-going CI/CT/CD pipeline to the incoming integration CI/CD/CT pipeline)
- Definition of content and format of all artifacts which are synchronized in both directions
- Automatic creation of optimised test plans based the incoming information from the collaborative "just-in-time" CI/CT/CD development tool
- Automatic generation of stacks adapted to the test environments
- Automatic creation of test environments

Validation & verification of ADAS/AD SW stacks

In the ADAS/AD domain, there's a significant lack of cost-effective, widely accepted verification and validation (V&V) methods and tools. Researchers like Winner et al. estimate that over 400 million kilometers of road driving would be necessary to statistically prove that an automated vehicle is as safe as a manually driven one. This makes traditional certification through physical road tests impractical.

To address this challenge, the development and use of Digital Twins has emerged as a crucial approach for complex automated systems, especially those interacting with the cloud. Meanwhile, experts worldwide some also under the guidance of UN/ECE, are working on standards for approving ADAS and AD functions. The draft of Regulation (EU) 2019/2144 for ADAS functions has been approved and includes definitions for expected functionality, along with initial guidelines on how to demonstrate their effectiveness.

The challenge lies in the intricate interaction between safety-critical automated systems and their environment. As automation levels increase, verifying not only the system's core functionality but also its responses to various scenarios becomes complex. These scenarios include both common situations (e.g., vehicles, trucks, airplanes) and rare safety-critical events, which can happen at any time at any scenario (e.g, sudden crossing of an animal, sudden blinding sun, sudden hail, …). Incorporating embedding machine learning (ML) and AI based algorithms and models introduces new additional challenges for verification and validation, necessitating novel techniques to ensure stability and safety while leveraging learning from new data. Updating existing systems without unintended consequences is crucial and requires new methods in efficient and fast testing of partially modified ADAS or AD systems.

The test areas include safe handling of internal failures as described in ISO 26262, safety of the system in the indented use (as described in SOTIF / IOS 21448, cybersecurity as well as the perceived value of the customer (see Figure 3.1-11)

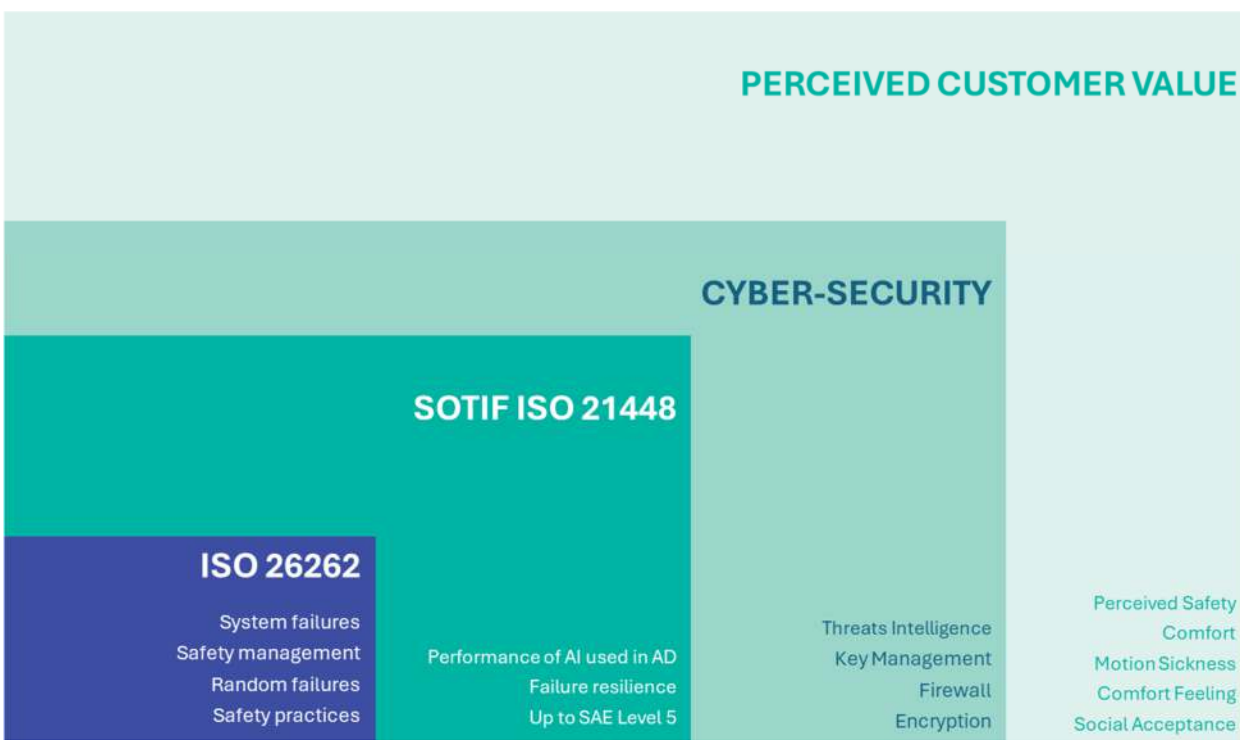


Figure 3.1-11 Validation framework for automated vehicles (Source: AVL List GmbH)

Modern highly automated cyber-physical systems increasingly dynamically evolve after their deployment, and OTA updates and upgrades are becoming necessary in such systems. This requires a tight integration of the analysis of data from the field and fast software changes with subsequent over-the-air updates of the vehicle in the field and ensure that any safety and/or security relevant problems identified in vehicles in the field are solved as soon as possible at all vehicles in the field. This requires new methods and trustable tools for these updates and upgrades, together with the respective re-verification and re-certification approaches, necessary to avoid negative impacts on both safety and security. These verification and validation tests must be integrated in fully automated CI/CD pipelines.

Automation functions of vehicles rely on environment sensors, such as cameras, lidar, radar and ultrasonic sensors, as well as communication to other vehicles or infrastructures. As these safety- relevant components may degrade over time or be exposed to cyber-threats, accelerated reliability, resilience and cybersecurity test methods are required. This will need further diagnostic devices to check the reliability of hardware, sensors, and their software. The role of the driver and any additional passengers in an automated vehicle is completely changing, and therefore new test methods and tools are necessary to ensure comfort and perceived safety (societal acceptance). These are already in the early development phases in terms of new functionality and their safety.

Many of the above issues are mentioned in existing or upcoming automotive standards for cyber-physical systems – for example, Safety of Intended Functionality (SOTIF), ISO 26262 and UL4600 (see Fig. 7). As none of these standards are mature enough to certify fully automated vehicles with reasonable effort, close cooperation between the standardisation committees and the research consortia will be necessary.

The expected outcome is twofold:

- Digital innovation to increase road safety as specified in the CCAM programme: reduce the number of road fatalities and accidents caused by human errors to zero by 2050, as well as ensuring that no added road fatalities are introduced by automated transport.
- Reduce validation costs down from the current two-to-five times of the implementation and testing of automation functions in mobility by 60–80%.

Key focus areas

To ensure the safety, security and comfort of automated mobility systems (including vehicles, ships, trains, airplanes, and off-road vehicles), verification, validation, and certification toolchains are essential. These toolchains address the test of components like environment sensors, communication systems, perception systems, route planning, and actuator systems. They also validate complete automated vehicles, assess reliability, and ensure secure OTA updates. Additionally, Digital Twins play a crucial role in validating new concepts and Car-2-Cloud interactions

Validation toolchains, their components and underlying methods should lead to safe, reliable, and secure argumentation, describing why the performed tests resulted in the estimated residual risk for automated cyber-physical systems for on-road or off-road vehicles, ships, trains, and airplanes. Optimisation methods can be used to balance multiple design objectives – e.g. ensure that the residual risk remains below a certain limit (such as stipulated by regulatory bodies) while meeting financial design targets. The verification, validation and certification tools and methods may be used for cyber-physical systems with different levels of automation.

Virtual validation, or more concretely scenario-based virtual validation, is considered a cornerstone for the verification, validation, and certification of vehicles. Two aspects are essential here: (i) scenarios representing the most relevant situations; and (ii) reliable simulation models. Scenarios may be derived from requirements of safety analyses, extracted from naturalistic driving, or synthetically created using gaming theory-based methods with a defined relevance. Statistical safety evidence from scenario-based verification and validation derived from naturalistic driving is needed. Also, here the establishment for open platforms and ecosystems for the creation and maintenance of reliable scenarios is encouraged. The definition of performance (safety, security, reliability, and comfort) indicators for different automation functions and SAE levels (in the case of road vehicles) and ODDs (Operational Design Domains) is necessary. Again, ecosystems to share these data are useful.

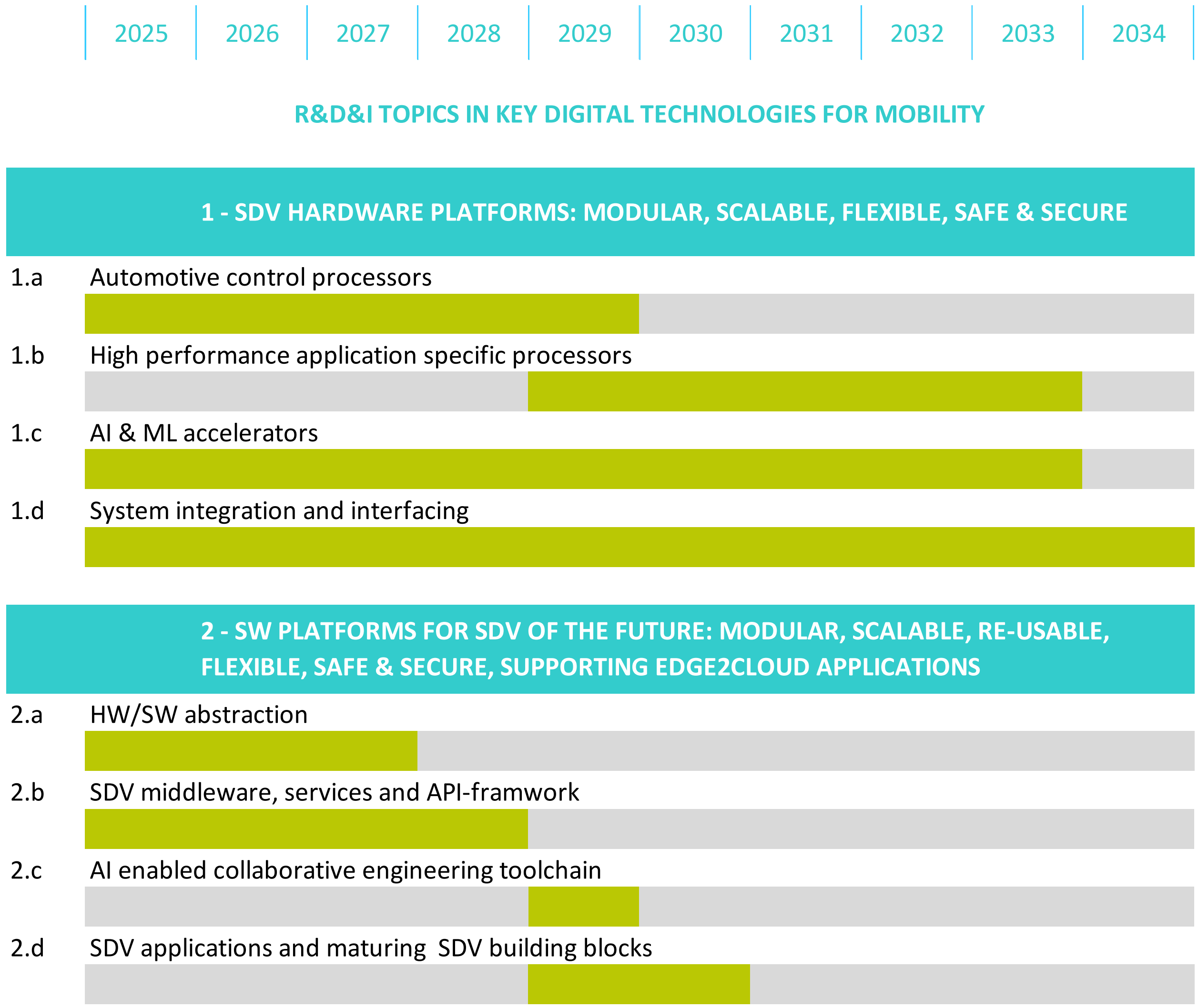
Reliable simulation models for environmental sensors, vehicles, drivers and traffic participants, as well as traffic, are vital. The development of these models, and the corresponding test systems, are essential. To test safety-critical scenarios using real vehicles in a safe environment requires the creation of stimulators for the different environmental sensors under different weather, traffic and road conditions. The verification, validation and certification of vehicles will be carried out with a combination of virtual test environments using model-in-the-loop (MIL) and software-in-the-loop (SIL) in the cloud with massive parallel processing in order to allow for testing of very high numbers of scenarios in combination with different critical events and varying ODD conditions as sun, rain, fog, snow etc., mixed virtual/real environments (vehicle-in-the-loop (VIL), and hardware-in-the-loop (HIL)), as well as a proving ground for real-world public road testing. Road testing will result in amounts of data larger than 20 TB per hour per vehicle, and therefore adequate data acquisition, management and (cloud or on-premises) evaluation systems capable of handling the specific data types of the sensors are critical (although these do not exist yet). Additionally, OTA data collection from in-use operations is required to continuously collect unknown scenarios that can be fed back into development to improve the quality of the systems.

Additional challenges covering this topic can also be found in Chapter 2.3 (Architecture and Design: Methods and Tools) and Chapter 2.4 (Quality, Reliability, Safety and Cybersecurity) of this SRIA.

3.1.3 TIMELINE

The roadmap for the key digital technologies in mobility are aligned with European roadmaps for terrestrial, water and aerospace transport:

- A new European Road Transport Research Advisory Council (ERTRAC) roadmap entitled "Sustainable Energies and Powertrains for Road Transport – Towards Electrification and other Renewable Energy Carriers".
- Urban mobility roadmap.
- Long-distance freight transport roadmap.
- Towards zero logistics emissions by 2050.
- The joint European Technology Platform (ETP) common paper published in 2019.
- The European roadmap on connected and automated driving published in 2019.
- The Joint Strategic Research Innovation and Deployment Agenda (SRIDA) for the AI, Data and Robotics Partnership (euROBOTICS), September 2020.
- 2Zero SRIA.


2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
R&D&I TOPICS IN KEY DIGITAL TECHNOLOGIES FOR MOBILITY
1 - SDV HARDWARE PLATFORMS: MODULAR, SCALABLE, FLEXIBLE, SAFE & SECURE
1.a Automotive control processors
1.b High performance application specific processors
1.c AI & ML accelerators
1.d System integration and interfacing
2 - SW PLATFORMS FOR SDV OF THE FUTURE: MODULAR, SCALABLE, RE-USABLE, FLEXIBLE, SAFE & SECURE, SUPPORTING EDGE2CLOUD APPLICATIONS
2.a HW/SW abstraction
2.b SDV middleware, services and API-framwork
2.c AI enabled collaborative engineering toolchain
2.d SDV applications and maturing SDV building blocks

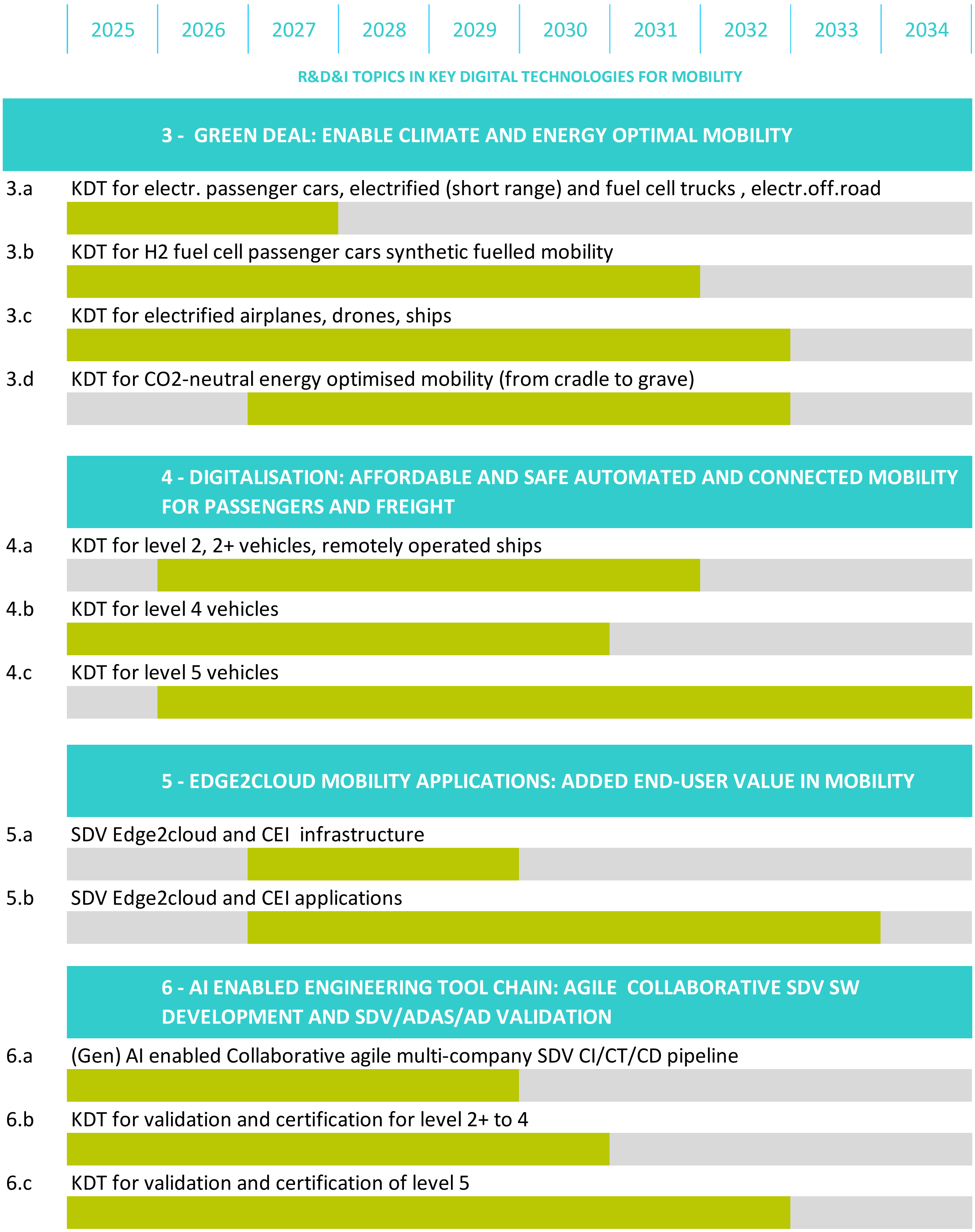


Figure 3.1-12 Roadmap for electronics and software based mobility applications

The roadmap combines the objectives in the application research programmes 2Zero and CCAM with the derived ECS mobility challenges. The following roadmap shows when R&D&I activities are required to ensure the key digital technologies are available for use in the different mobility domains. In areas that already include ongoing electric mobility, the focus is more on improvements to existing concepts (for example, optimisation of costs), while for others (such as electric aircrafts) it is more about focusing on lower technology readiness levels (TRLs). These are both going on in parallel and are also influencing each other. This roadmap is a preliminary estimate when the ECS will need to be ready for the various technology fields. It will be annual updated as new domain roadmaps become available.

3.2

ECS Key Application Areas

ENERGY

3.2 Energy

3.2.1 Scope

Change towards the carbon neutral society and challenges for ECS

Energy systems *supplying clean, affordable and secure energy* are the focus of *The European Green Deal*. To achieve this goal, the European Union set targets for a renewable energy share of 32 percent and a Greenhouse gas emission reduction of 55 percent by 2030. Renewable energies bring several benefits such as mitigating climate change, emission reduction as well as improvements in the European energy security1. Although the EU even surpassed its' 2020 target of 20 percent, sustained action with an accelerated pace is necessary to prepare the economy and society for the upcoming climate challenges. The drop in CO_2 emissions was overcome rapidly by the uneven recovery from the Covid-induced recession (Figure 3.2.1), which put a major strain on the European energy system with a rebound in coal and oil use2.The power sector must be further transformed from fossil fuel-based to renewable generation and, at the same time, needs to grow in order to enable decarbonisation of mobility, industry, and thermal energy supply, and reach the climate targets. The recent change in the supply strategy in Europe to be independent of strategically critical gas or oil suppliers is a further boost for renewable energies and efficiency measures. The shortage of materials (e.g. batteries and other electronic equipment) has already had a serious effect on the R&D and direction of developments. Some materials are getting rarer or can contribute to conflicts. Furthermore, there is a growing shortage of skilled workers, which is a huge societal challenge and needs to be compensated by fast technological progress and innovation. Therefore, smarter components are needed to compensate the growing shortage of technical knowledge and skills.

Because of the increasing residual load, resulting from the local mismatch between decentralised renewable generation and load, a digitally controlled transmission and distribution infrastructure is required. Thus, electronic components and systems (ECS) are key to future energy systems being optimised in both design and operation, for high efficiency, low CO_2-emissions, cost, and security of supply. The development of energy systems is driven by action against climate change, booming decentralised renewable generation (solar, wind), digitalisation and AI technologies, as well as cyber security issues. The Energy Chapter highlights the Major Challenges in the changing energy landscape based on electrical energy generation, supply, conversion, and use. Highest efficiencies and highly reliable, secure solutions are required to achieve the change towards a carbon neutral society in 2050.

---

1 European Environment Agency (2022). Share of energy consumption from renewable sources in Europe. https://www.eea.europa.eu/ims/share-of-energy-consumption-from

2 International Energy Agency (2021). World Energy Outlook 2021.

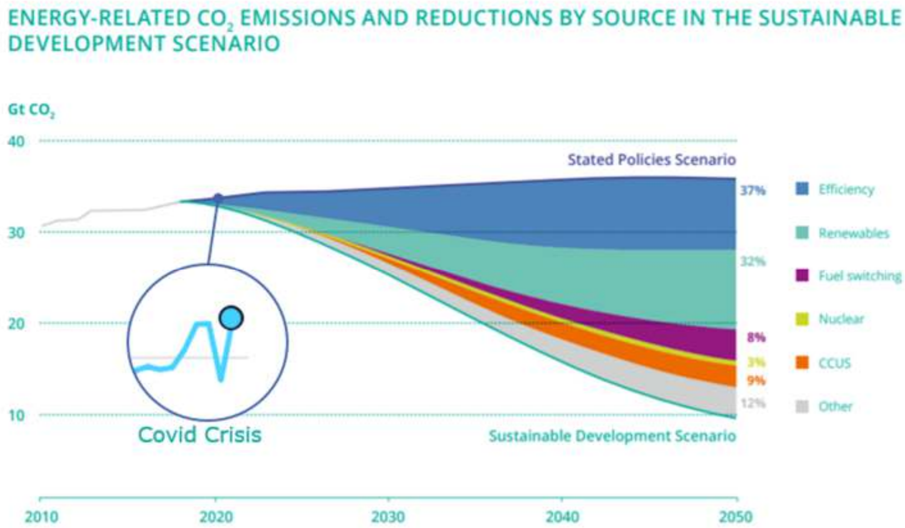
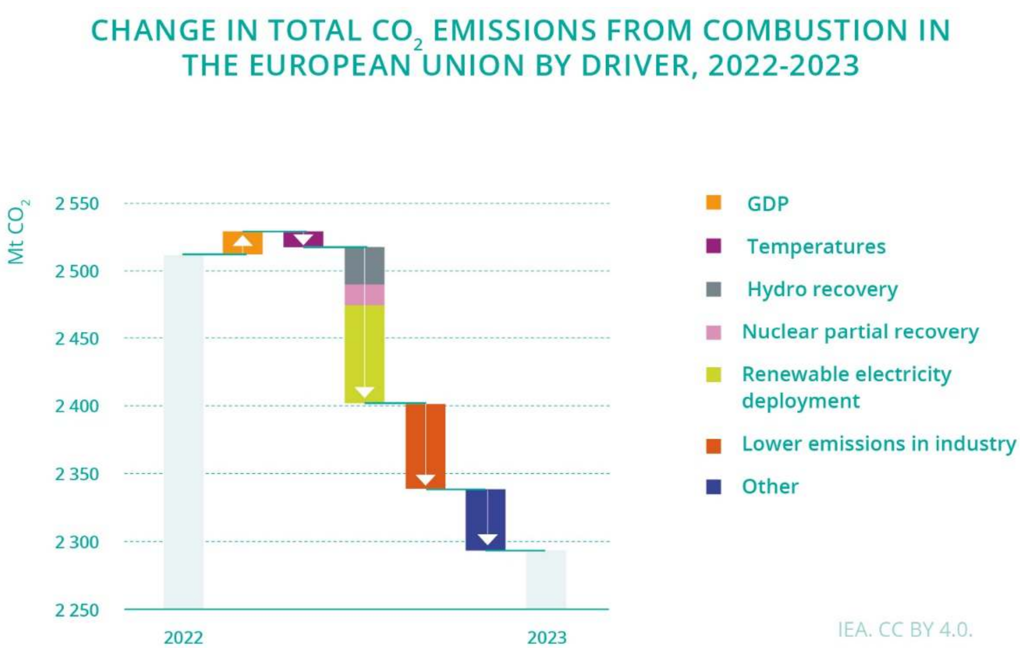


Figure 3.2.1 - Efficiency and renewables provide most potential for CO_2 emissions reductions. Source: IEA World Energy Outlook 2019. In the graph the impact of Covid in 2020 & 2021 is indicated, emission level is back on the levels from before the Covid crisis. Source: IEA Global Energy Review: CO_2 Emissions in 2021.

According to the IEA "CO_2 Emissions in 2023" report for the European Union the trend goes towards less emissions, main factors are an increased renewable electricity deployment and lower emissions in industry. The challenge is to accelerate that trend and manage the challenges coming up by local use and balance to integrate the decentralized highly variable supply into the grid:



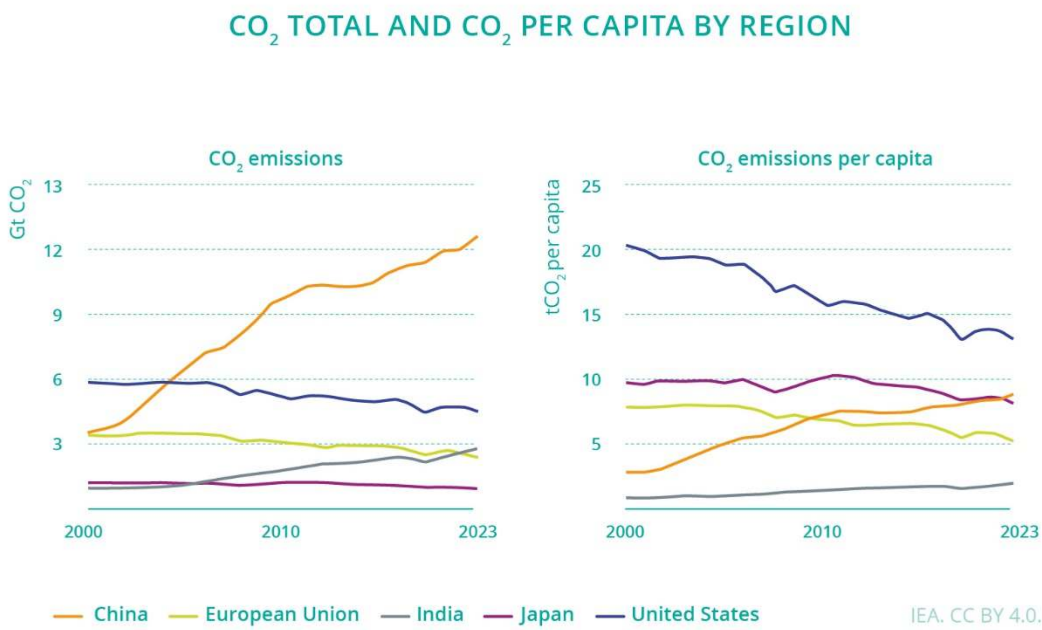


Figure 3.2.2 & 3.2.3 - Change in total CO_2 emissions in the EU & CO_2 emission globally – clear trend to less emissions and dominating factors relevant for less emissions in the EU visible. Source: IEA – CO_2 emissions in 2023: A new record high, but is there light at the end of the tunnel?

3.2.2 Application trends and societal benefits

Application trends

At present, 75 percent of total greenhouse gas emissions in the EU come from the energy sector. The energy world is undergoing a radical transformation: promoted e.g. by EU and national roadmaps, the globally installed capacity of renewable generation has doubled within the past 10 years. Europe alone expanded its renewable generation capacity by 6.4% in 2021. This increase is dominated by wind and solar energy being characterised by strongly intermittent, distributed generation. Altogether wind and solar energy made up one fifth of Europe's electricity generation in 2021 with plant capacities ranging from domestic solar (≤ 10 kW) via commercial solar and wind (≤ 500 kW) to power stations at utility scale (≥ 1 MW). At the same time, the levelised cost of electricity (LCOE) from photovoltaic (PV) sources dropped by 13 to 15%. However, the rise of renewables is still too slow - wind and solar generation growth must nearly triple to reach Europe's 2030 green deal target. In the long term, it enables the substitution of fossil fuel-based transportation, domestic heating, and commercial & industrial processes as well as address the strong economic growth of non-OECD countries. Since the pursuit of all economically viable opportunities for efficiency improvement can reduce global energy intensity by more than 3% each year, increasing energy efficiency may be accountable for 30 % CO_2 emission-reduction by 2050 with current policy settings, but can be even increased up to 40 % if worldwide announced climate pledges are met. Energy supply to all sectors affordably and reliably needs to match the demand and availability as efficient as possible (Figure 3.2.4).

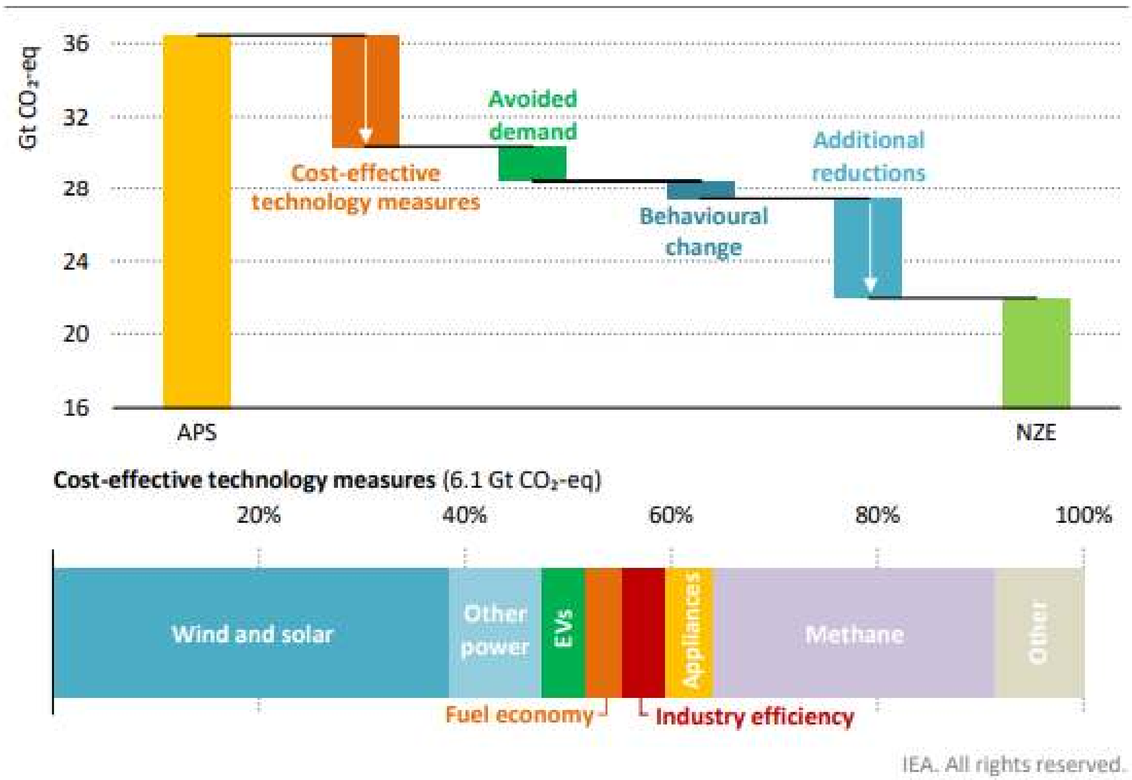


Figure 3.2.4 - Measures to reach the Net Zero pledge announced for 2050. Source: IEA World Energy Outlook 2021.

Thus, the power grid architecture developed for centralised, unidirectional, demand driven power generation will be transformed into a multi-modal energy system (MM-ENS) architecture (Figure 3.2.5). It will comprise distributed renewable generation, energy conversion units for sector coupling, transmission and distribution grids allowing bi-directional power flow, and energy storage for all modes of energy (electric, thermal, chemical). Energy management systems (EMS) will optimise ENS-operation. It will match load and demand at all levels ranging from the nanogrid (behind the meter, building level) and the microgrid (district or community level) to the regional distribution grid, which is connected to the cross-regional transmission infrastructure. Fossil-fuelled power plants, which used to operate on schedules orienting at the demand, will turn into back-up power supply facilities.

The overall reduction of energy consumption in addition to efficiency measures will be always a target, since all energy usage that can be avoided also implies reduction of emissions. This can be achieved by control elements for switching off energy use and zero power stand-by functionality or by transformation to new technologies as in the last decade the transfer to LED illumination had a high impact. Upcoming threats are energy consuming ICT technology related applications like blockchain, AI, data traffic, or digital currencies. The challenge will be to develop highly efficient algorithms and methodologies to decrease energy consumption despite the increased use of these new technologies. Between 2012 and 2018, the amount of computing power required for cutting-edge large AI models doubled every 3.4 months, amounting to a more than 300,000-fold increase [3] and this trend will hold until more research goes in efficiency. Quantum computing have in general the potential to decrease energy consumption for solving specific problems additionally. Already within the design and development cycle sustainability goals, energy efficiency and environmental legislation will need to be taken into account.

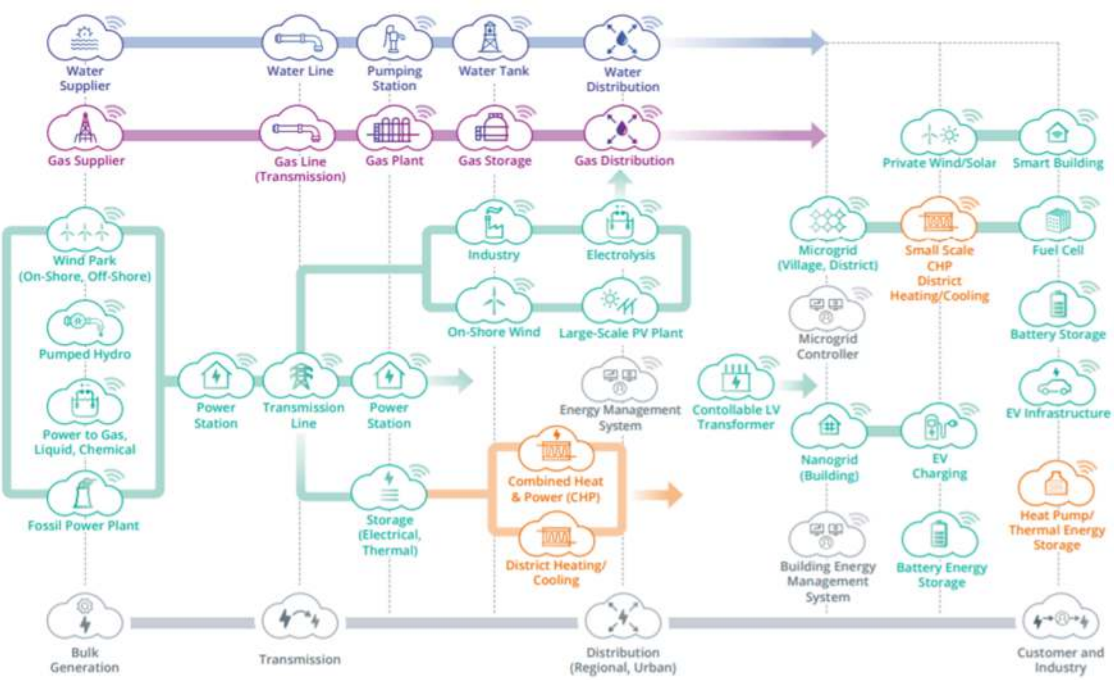


Figure 3.2.5- Interconnected Energy Infrastructure; Source: Siemens Corporate Technology

Key to these new energy applications will be smart sensors, networks of sensors, and smart actuators that enable status monitoring on each grid level as well as smart converters (for all voltage levels). The converters need to use highly efficient and fast semiconductor power devices and modules that enable real-time control of energy system components and grids for optimised operation based on forecasts of generation and demand but also in case of any critical

[3] "Are quantum computers really energy efficient?", Sophia Chen, Nature Computational Science volume 3, pages457–460 (2023)

event. The future grid operation requires a sophisticated information and communication infrastructure including cloud services, IT security, and AI technologies. Altogether, they will contribute to significant reduction of energy consumption and, consequently, CO_2 emission.

To achieve the targets of the Green Deal and to have competitive advantages for European based technologies and solutions, research has to be performed in the following areas:

(1) Significant reduction and recovery of losses (application and SoA-related).

(2) Increase of power density and reduction of losses (e.g. through exploitation of new materials) and a decrease of system size by miniaturisation and smartintegration, on the system and power electronics level.

(3) Increased functionality, reliability, and lifetime (incl. sensors & actuators, ECS HW/SW, semiconductor power devices, artificial intelligence, machine learning, monitoring systems, etc.).

(4) Manufacturing and supply of energy relevant components, modules, and systems.

(5) Management of renewables via intermediate storage, smart control systems, share of renewable energies, peak control or viability management for the increase of energy flexibility. Grid stabilisation through e-vehicle charging.

(6) Energy supply infrastructure for e-mobility, digital live, and industry 4.0.

(7) “Plug and play integration” of ECS into self-organised grids and multi-modal systems, real-time digital twin capability in component and complete system design (to simulate system behaviour).

(8) Safety and security issues of self-organised grids and multi-modal systems through smart edge devices and high-level IT security (resilient communications and trustworthy AI).

(9) ECS for energy storage technologies: production, transportation, storage, distribution, combustion and energy conversion systems.

(10) Optimisation of applications and exploitation of achieved technology advances in all areas where electrical energy is consumed.

(11) Energy technologies in the circular economy approach: predictive and condition-based maintenance with repair, refurbish, reuse and recycle capabilities, LCA of ECS and reduction of environmental impact

(12) Aligning with standardisation of our energy systems.

(13) Manufacturing and world-leading technologies for energy relevant applications in Europe.

(14) Scheduling for cost-efficient energy consumption.

(15) Involvement of the consumer: traceable eco-footprint and incentives towards environmentally-friendly behavioural change.

(16) Design and development strategies of ECS that optimise the total environmental impact of energy solutions (e.g. trade-offs between environmental footprint and handprint, rebound effects)

(17) Innovation strategies for regulations-enabled energy markets and technologies.

External requirements and Societal Benefits

In alignment with the **Parisian Agreements**, the EU committed to substantial reductions of CO_2 emission. In particular, the EU aims to make Europe the first climate-neutral continent by 2050 (EU long-term strategy) while boosting the competitiveness of the European industry. Carbon pricing throughout the EU economy is going to be implemented more strictly. Further climate laws will be introduced and continuing policies will be clarified by the European Commission in 2022. The new policy regarding **"Clean energy for all Europeans package"** was completed by the EU in 2019 as a comprehensive update of its energy policy framework and updated with the new Green Deal in July 2021. It emphasises renewable energy, energy performance of buildings, energy efficiency, governance regulation, and electricity market design. Smarter buildings with more automation and control systems for effective operation shall be promoted. E-mobility infrastructure is going to be supported further. Energy efficiency targets and energy labels were tightened to encourage the industry to innovate.

Figure 3.2.6 - Energy from renewable sources: Wind turbines and photovoltaic (Source: © Mariana Proenca/Karsten Wurth – Unsplash)

To achieve the **Green Deal** goal of "clean, affordable and secure energy" in all sectors, new laws and regulations will be required. While subsidies and regulations will promote sustainable developments in all application domain of ECS (energy, industry, mobility, communication, consumer goods, and cities), the energy domain with targeted 40% renewables in the energy mix until 2030 is the foundation to all of them. Additional perspectives are given by the United Nation's "**Roadmap 2050**" addressing sustainable development solutions and implementations towards a carbon-neutral global population.

Figure 3.2.7 - Electrification of the transport sector.

All these factors are considered for the roadmaps on research, development, and innovation of ECS for the applications in the energy sector. Potential targets comprise the implementation of electricity storage solutions (e.g. vehicle2grid, battery grid storage), the further increase in efficiency and the reduction in life cycle costs of energy generation from renewable sources (Figure 3.2.6), the electrification of transportation (Figure 3.2.7), and the thermal processes in industry as well as the development of secure, self-learning energy management systems for buildings and industrial sites. ECS as enablers support the EU and national energy targets to achieve sustainability (Figure 3.2.8) and are essential for a highly developed energy landscape towards a fair, democratic, healthy and prosperous society.

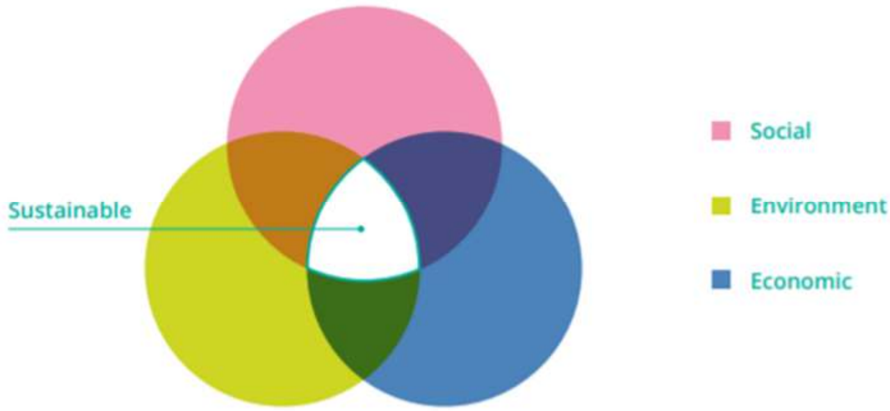


Figure 3.2.8- Three pillars of sustainability (Source: Purvis, Mao, Robinson 2018: Three pillars of sustainability: in search of conceptual origins).

Energy efficiency through ECS fosters economic development towards a circular economy and new employment opportunities. They will have a huge impact on job generation and education if based on the complete supply chain and fully developed in Europe. With more than 11 million jobs in the field of renewable energies4 and indirectly involved technologies, this is a visible and significant factor for economic and societal stability. The capability of maintaining the understanding of the complete systems as well as the competence from small-scale solutions up to balanced regional energy supply solutions are key to the European competitiveness and success in the global market of energy solutions. Also the consumer itself can contribute its share, thus consumer empowerment to energy savings and efficiency should be taken into account for the development of energy systems. Societal benefits include access to knowledge, development of modern lifestyle and the availability of energy all the time and everywhere – with a minimum of wasted energy and a minimum of greenhouse gas emissions. Therefore, ECS and its application domains enable Europe to meet the needs of the present without compromising the ability of future generations to meet their own needs.

3.2.3 Major Challenges

Five Major Challenges have been identified for the energy domain:

- **Major Challenge 1**: Smart & Efficient - Managing Energy Generation, Conversion, and Storage Systems.
- **Major Challenge 2**: Energy Management from On-Site to Distribution Systems.

---

4 IRENA (2019), Renewable Energy and Jobs – Annual Review.

- **Major Challenge 3**: Future Transmission Grids.
- **Major Challenge 4**: Achieving Clean, Efficient & Resilient Urban/ Regional Energy Supply.
- **Major Challenge 5**: Cross-Sectional Tasks for Energy System Monitoring & Control.

3.2.3.1 *Major Challenge 1: Smart & Efficient - Managing Energy Generation, Conversion, and Storage Systems*

3.2.3.1.1 Status, vision and selected outcome

According to the IEA's Efficient World Strategy, digitalisation enhances energy efficiency gains in the transportation and industry sectors[5]. Smart and efficient energy systems are drivers of energy savings. Therefore, they are in full alignment with the Green Deal. Alternative ways of energy generation (hydro, photovoltaic, and wind) and the electrification within the industry, the transport / mobility, and the construction / building sectors result in the challenge of creating smart, efficient, and reliable energy generation, conversion, and storage components.

Smart Energy Systems

For operating smart energy systems, all the energy conversion and storage components need to be equipped with smart actuators and sensors for status and health monitoring as well as optimisation of grid operation. The integration of sensor, connectivity and edge processing in supplementary/additional parts will enable the creation of intelligent facilities by retrofitting. The creation of secure electronic control units requires development of specific hardware and software. Scalable modular renewable energy supply with seamless installation capability brings in or scales up renewables from single households to larger installations.

Consequently, smart control units need to be developed for all types of energy production, conversion, and storage components comprising smart electronic converters, actuators, sensors, security systems and reference communication interfaces. They shall have plug-and-play functionality and real-time digital twin capabilities in component and complete system design to simulate system behavior for evaluation of its' health status.

For offshore energy generation, such as windfarms and tidal energy generators, fibre optical sensors is an emerging technology beneficial for online monitoring of metal fatigue and excessive turbulences. This technology is currently being developed for such monitoring in aircraft wings and ship masts.

Conversion

Electrification of industry is one of the main implications to reach the 2050 decarbonisation targets, mainly via the conversion from fuel-based heating processes to electro-heating solutions.

5 IEA (2019), Energy Efficiency.

In addition, direct electrification of industrial production processes (such as electro-synthesis of chemicals or electrolysis) is also crucial for replacing present CO_2 emitting solutions. In the case of Heating, Ventilation, and Air Conditioning (HVAC) systems, significant reductions in consumption can be obtained by optimizing the system that handles all the processes of energy management or by changing the use of the Machine-to-Machine (M2M) technologies. For both strategies, efficient ECS are required to obtain optimal control functionality based on sensing, collecting, processing, and evaluating device related data. DC power supply requirements based on advanced semiconductor power devices will provide lower power consumption and thus, feature higher efficiency of the increasing ICT energy consumption (i.e. through data centres) . Investments in the next-generation computing, storage, and heat removal technologies will be required to avoid a steep increase on energy demands and to minimise the implications of unavoidable data centre energy use on the global climate. In data centres and 5G/6G networks, photonic ICs can route information streams from fibre to fibre without conversions into electronics in between. This will be highly efficient and save energy. The advanced features of 5G/6G will innovate the use of the technology (Figure 3.2.9), but as consequence of larger data rates and through-puts, cost and energy demand will increase substantially. Therefore, energy harvesting capability of sensors and devices in the 6G environment will be one of the crucial aspects towards a green and cost-efficient technology landscape[6].

Figure 3.2.9 - 5G as enabler of an interconnected smart network. Source: European Commission, Towards 5G.

Power electronics circuits based on semiconductor power devices are used in all conversion processes. Silicon based power devices are approaching their ultimate limits in terms of breakdown voltage, current, switching frequency and temperature capabilities. Next generation power

6 Routray (2016), Green initiatives in 5G, 2nd International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), Chennai: 617-621.

semiconductor devices will rely on Wide Band Gap (SiC, GaN) and Ultra WBG (diamond, Ga_2O_3) technologies. The integration of ultra-wideband gap (UWBG) semiconductors presents a transformative opportunity for power electronics. With thermal conductivities far surpassing those of traditional materials, UWBG semiconductors such as diamond offer unparalleled heat dissipation, enabling the creation of more compact, efficient, and high-performing power modules. Their capacity to handle extreme voltages further enhances device robustness and energy efficiency. As a strategic initiative, prioritizing research and development in UWBG materials could catalyze breakthroughs in power electronics, paving the way for advanced technological solutions that align with the goals of miniaturization, efficiency, and long-term reliability.

Due to this unstoppable trend, research on device reliability, packaging and assembling methods suitable for very high electric fields and high temperature, is strongly required. A focus also needs to be set on the medium voltage grid (< 45 kV) , as the power rating of applications will increase beyond 1 MW (EV charging, BESS, hydrogen electrolyzers etc.) and will be preferably connected directly to the medium voltage. Here, the case of solid-state transformers based on medium frequency transformers will become a cost-effective alternative in the near future and consequently help to reduce the material spending for standard grid frequency transformers. As many concepts rely on a modular approach using power electronics building blocks (PEBB) aspects of coordinated system controls, reliability based on redundancy need also to be investigated.

Storage

Energy storage deployment provides energy system flexibility. Looking at further storage possibilities, different options for various capabilities need further efficiency improvements. As an example, optimised converters, sensor solutions for monitoring, and battery management systems need to be developed for storage options, all including ECS. In power generation, hydrogen with its many uses (Figure 3.2.10) is one of the leading options for storing renewable energy. Hydrogen can be used in gas turbines to increase power system flexibility. In combination with fuel cells, it is also a great vector of clean energy since it allows to produce electricity directly onboard of EV or in areas, which are cut off from the power grid. With declining costs for renewable electricity, interest is growing in electrolytic hydrogen. ECS will be employed in electronics for electrolyzers, fuel cells, as well as power management and health monitoring.

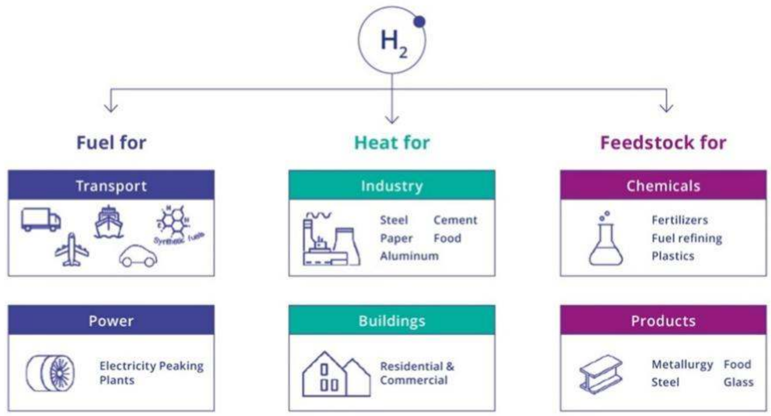


Figure 3.2.10 - The many uses of Hydrogen, source: Bloomberg NEF.

3.2.3.1.2 Key focus areas for increased efficiency and smart energy generation, conversion and storage components

- Increased efficiency at all levels:
  - Power conversion and wide-bandgap semiconductor power devices.
  - Power supply.
  - Energy harvesting.
  - Energy management.

- Residential, commercial, and industrial demand side management (scheduling and load adaption):
  - Sensors, actuators, drives, controls and innovative components.
  - Full monitoring in adaptive and controlled systems.
  - High efficiency electric drives, heat pumps, cooling, HVAC, data centres and other consumers of electricity for variable load operation.
  - Solutions for increasing power demand of 5G/6G systems.

- Development of Energy Management Systems including:
  - Optimisation module.
  - Demand and generation forecast.
  - Customer preferences.
  - Weather forecasts.
  - Price/tariff information/forecast for scheduling controllable loads and generators.

- Smart sensor network: internal and external physical parameters that influence energy conversion efficiency.
- Resilient and smart communication and edge devices.
- Deployment of Trustworthy AI.
- Fiberoptic sensors for fatigue detection.
- Converters for power quality improvement (e.g. electronics filters to manage resonances).
- Sensors and controls for the management of decompression and compression and leakage detectors for methane, hydrogen, and other gases.
- ECS for the coupling of processes in the chemical and electrical industry.
- Traceability and labelling of green energy.

- Conversion of industrial processes:
    - "Industrial electrification" (Replacement of CO_2-emitting processes by others based on "clean" electricity).
    - Electric drives for commercial & industry applications.
    - Industry 4.0 with combination of Cyber-Physical Systems (CPS), Internet of Things (IoT), Artificial Intelligence (AI).
    - DC subsystems for industrial production / data centre applications and DC distribution grids.
    - Photonic routing in data centres from fibre to fibre without conversion to electronics.
    - Carbon capture technologies compensating production emissions (up to negative emissions).

- Development and application of storage optimised for residential, commercial, industrial utilisation:
    - Control, interfaces to batteries, fuel cells, hydrogen storage electrolysers.
    - Integrated battery driven applications (e-car charging, PV – system local storage).
    - Power Storage to "buffer" net fluctuations and to avoid long distance transmission.
    - Smart storage technologies from low to medium voltage.

3.2.3.2 Major Challenge 2: Energy Management from On-Site to Distribution Systems

3.2.3.2.1 Status, vision and expected outcome

The distribution grid comprises commercial scale renewable generation as well as private smaller renewable power generation units, conversion between different energy modes, storage, control and protection systems for the grid infrastructure together with all kind of consumption.

Autonomous Control Systems

In the future distribution grid, generation and consumption by power electronics systems will surpass the share of synchronous generation. This leads to potential grid instabilities due to lack of inertia. Therefore, autonomous control systems need to be implemented to control the high demand loads. These control systems should be organised hierarchically to adjust the heavy loads according to the actual local production and storage capabilities so that import or export of power is minimised. Price control systems such as TOU (Time of use) can help to prevent grid violations. Storage devices, such as local community storages or e-Vehicles, can be charged when the price is low and discharged when the price is high, to provide flexibility as well as to ensure stability and reliability in the grids.

For industry or larger groups of buildings, control methods increase the flexibility of the total system and can be set up using hierarchical and intelligent control methods to minimise costs and to provide peak-shaving (Figure 3.2.11). For larger power production facilities, hybrid generation and storage solutions are also discussed, which integrate the power production facilities with storage devices to have the best arbitrage cost. Novel grid architectures for manufacturing strive to increase topological and energy flexibility within production cells to enable adaptive production optimisation. Also, blackouts and their consequences need to be prevented, since for example for large industrial electrolyzers they result in serious safety and cost issues.

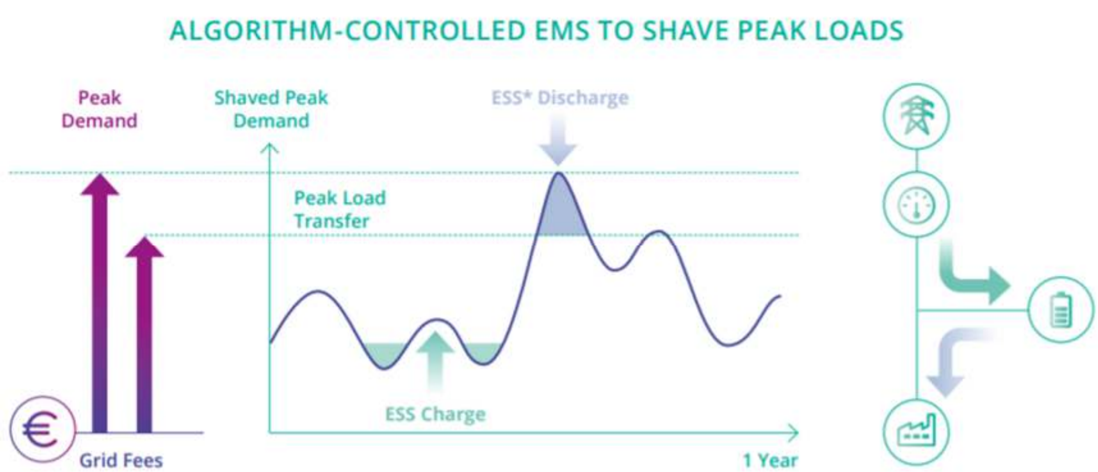


Figure 3.2.11: Visualisation of an algorithm-controlled energy management system to shave peak loads. An energy storage system predicts load peaks and charges/discharges a battery accordingly. Source: EDF Renewables.

Security, Reliability and Stability of Energy Systems

For stable, resilient on-site energy systems, multi-modal energy management systems allowing integration of electricity, heating & cooling, molecules, and transport (e-vehicle charging incl. Vehicle-2-Grid) will be developed. Their features comprise high level IT security, energy trading via local energy market platforms, renewable energy certification, development of solutions for low voltage electronic systems that are easy to setup, as well as support for self-learning against evolving needs.

Energy Management Systems for industrial and residential customers include optimisation module, demand and generation forecast, customer preferences, weather forecasts and price/tariff information/forecast. They require beyond-state-of-the-art techniques for scheduling controllable loads and generators, and to forecast the weather to produce accurate generation profiles. Furthermore, the interface to the grid might be used for additional power quality services based on power electronics converter technologies beyond state-of-the-art reactive power compensation (e.g. virtual inertia and balancing).

3.2.3.2.2 Key focus areas for on-site or behind the meter systems

- Security, reliability and stability of total energy system:
  - Automation of grids.
  - Storage of data.
  - Trustful AI and ML for optimised operation of the grid.
  - Machine-learning based forecasting algorithms for generating accurate generation profiles of expected power production and consumption.
  - Deployment of sensors and edge computing devices to health-check grid assets to increase lifetime and optimise operation.
  - Converters for power quality improvement (e.g. electronics filters to manage resonances).

- Stable and Resilient On-Site Energy Systems:
  - Integration of electricity, heating & cooling, molecules, and transport (e.g. Vehicle2X).
  - Coupling with energy trading systems, e.g. local energy market platforms.
  - High level IT security.
  - Renewable energy certification.

- Hybrid solutions:
  - Integrating power production facilities with storage devices.
  - Arbitrage cost, keeping level of production according to market bid.

- Virtual markets:
  - Flexibility in demand & supply.

- Aggregation of Energy consumption and production.

- Electric energy supply for manufacturing:
    - Higher uptime using novel industry grids and UPS.
    - Stable power supply using novel electronics converter technologies.
    - Blackout prevention.

- Plug-and-play capability for components, self-learning:
    - Integration of low voltage systems using flexible planning rules.
    - Cost effective solutions to minimise set up-time and manual parametrisation.
    - Reduced physical size and weight of individual transformer stations with equivalent power ratings.
    - Development of solid-state transformers with:
        - New functions for the operation of power systems.
        - Avoidance of infrastructure extensions caused by increasing share of distributed generation.

3.2.3.3 Major Challenge 3: Future Transmission Grids

3.2.3.3.1 Status, vision and expected outcome

New grid challenges

Future transmission and distribution grids will remain an integral backbone of energy systems. Coupling of different domains like electricity, thermal, gas etc. will enable new business opportunities which require new technological solutions for high power electronics, combined with sensors and ICT for monitoring, intensive control and prediction.

The energy generation and energy consumption pattern will drastically change as the industry and society at large will be highly electrified. Base industries such as the chemical industry, steel and cement production will completely change production technology to enable fossil free production and will require extreme amounts of electric energy. New industries such as giga volume battery production factories are planned in several places in Europe7. This together with a massive expansion of supercharger station for private passenger cars and heavy trucks with individual charge capacity of more than 1 MW will put severe challenges on the grid capacity in both networking and electronic components to manage highest possible efficiency. Therefore, further ECS R&D needs to work towards improvements of the grid capacity with the highest possible efficiency (Figure 3.2.12). Thus, continued development of components for HV transmission for 1.2 MV or even higher voltages are needed to roll out an efficient energy

7 Maisch, M. (2020). Europe’s Gigafactory Boom in Full Swing with Another Plant Announcement. PV Magazine.

transmission over Europe. In addition, new business models must be developed for the electric energy market enabled by the smart grid technology.

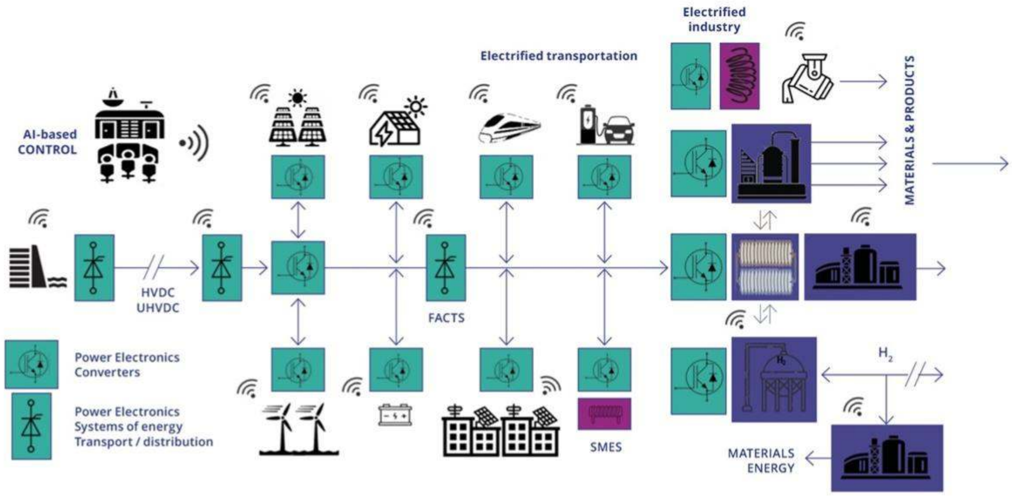


Figure 3.2.12: Power ECS at each point of the future transport and distribution grid. Source: CSIC Scientific Challenges: Towards 2030, Volume 8.

Resilience

To account for adverse conditions caused by climate change, the new national and transnational grids must include autonomous electricity generators based on fuel cells or local storage systems for communications and network information management as well as water-resistant components or modules. Also, sensors need to be placed at critical points to immediately alert authorities in case of unexpected incidents. To be able to quickly react to an electricity line fault, the system will benefit from powerful switches and AI to successfully reroute the systems. Additionally, predictive maintenance (e.g. with digital twins) of the energy supply sensors provides further safety and resilience. Due to the weak tectonic movement in most of the parts in Europe, transmission grids could become much more resilient and loss-less when buried in the ground. Thus, extra isolation technology needs to be considered and critical points equipped with smart ECS for monitoring, control, and prediction.

3.2.3.3.2 Key focus areas

- Grid stability during the industrial transition:
  - Efficiency increases.
  - Development of smart medium voltage grid.
  - Development of components for HV transmission for 1-2 MV or even higher voltages.
  - New solutions for high power electronics, combined with sensors and ICT for monitoring, control and prediction.
  - Development of new simulation and business models to foster innovations regarding grid stability.

- Development of a Trans-European energy infrastructure:
  - Secure, cross-regional transmission infrastructure.
  - Multi-terminal HVDC systems connecting remote energy generation sites.
  - Interaction between distribution systems on community and district level.
  - Development of components for HV transmission for > 1.2 MV.
  - Minimise Losses.
- Requirements on ECS by disruptive changes in transmission and use:
  - Flexibility in system design and operation.
  - Water-resistant components/modules.
  - Autonomous electricity generators based on fuel cells.
  - Modelling, sensing and forecasting weather conditions and thus, supply and demand.
  - Intelligent power devices, systems, and switches.
  - Status-/health-monitoring (e.g. ice sensor/detection) for transmission lines.
  - ECS for multi-modal energy systems.

3.2.3.4 Major Challenge 4: Achieving Clean, Efficient & Resilient Urban/Regional Energy Supply

3.2.3.4.1 Status, vision and expected outcome

A 40% renewable energy share in the electricity sector in Europe by 2030 needs additional decentralised, intermittent energy sources, bi-directional grid and storage for energy supply in transport, industrial and smart cities applications.

Multi-energy Systems (MES)

MES help to achieve optimised energy management. All sectors are integrated to maximise overall system efficiency. Energy flows between sectors and their storages ensure the highest use of renewable energy while balancing fluctuations.

Heating supply uses district heating, supported by heat-pumps and boilers, using thermal storage in the district heating system (Figure 3.2.13). Integration with industry makes use of waste process energy using heat pumps to boost from low (40-50 deg) to high temperatures in the pipe (80-90 deg). Electrolyzers add to the gas system or transport. Water treatment uses excess power from renewables adding further flexibility.

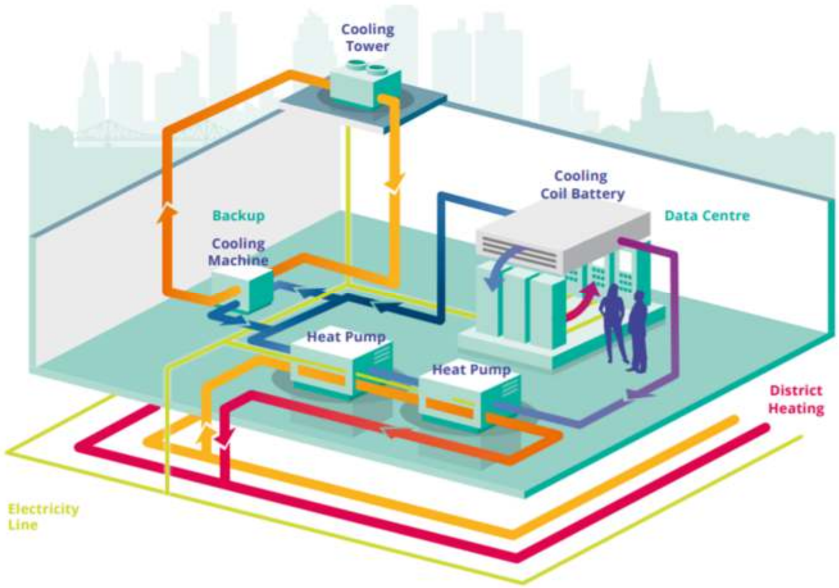


Figure 3.2.13: A Combination of heat pumps and district heating. Source: IEA HPT.

Local communities use MES concepts on regional level. Different local inputs are gathered for an overall aggregated control for the larger regions as well. Autonomous controllers are used behind the meters to support overall control. A clear hierarchical set up, control structure and knowledge of market interactions are necessary.

Complex integrated control systems use AI, machine learning and comprehensive communication grid/IoT platforms (including edge computing) to get all data for control and optimisation. Risk and security analysis provide resilience and ensure stability of MES.

Urban Transformation

Emission free cities use electrification and decentralised storages to improve efficiency and reliability. ECS as indispensable components ensure efficient management of data and data storage. AI approaches and the ECS supply chain for integrated applications in energy are key enablers for smart power grids. Electrification of urban mobility supports individual and public transport (incl. utility EV) and furthermore, contributes to the stabilisation of the grid. The first needs household and public charging, the latter uses well defined charging points on (bus) lines or at terminals. Powers vary from 10 kW (LV) to 600 kW (MV). Reservation and optimisation are based on ICT.

Other, crucial aspects of emission free cities are an efficient urban energy infrastructure, low carbon and smart residential and service buildings, low carbon mobility, smart water systems and smart waste management. Even the shift to LEDs without any smart functions can result in energy savings of ~50% in an industrial setting[8]. Carbon capture technologies will add another dimension to the energy systems.

---

8 Muneeb A, Ijaz S, Khalid S, Mughal A (2017) Research Study on Gained Energy Efficiency in a Commercial Setup by Replacing Conventional Lights with Modern Energy Saving Lights. J Archit Eng Tech 6: 202.

Storage Solutions

In households, battery energy storage devices can used to increase self-consumption. Some regions will use heat/cooling storage. Algorithms/models for optimal use of storage (community/private/ industrial) are based on technical parameters, demand and generation forecasts, customer preferences, in order to reduce power peaks and to support integration of RES into existing infrastructure.
MES in larger communities with different kinds of storage possibilities (electrical, thermal, gas, water etc.) play an important role. V2X is used as huge distributed electrical energy storage. Systems with electrolyzers might use storage tanks for gas production. Thus, development of grid-supporting control algorithms and supporting regional energy management for communities (e.g. P2P trading via storage systems, self-consumption optimisation) are needed.

3.2.3.4.2 Key focus areas for achieving efficient community and regional energy management

- Electric Energy Supply for urban mobility:
  - Development of household and public charging infrastructure.
  - Creation of HV (wireless) charging points along the (bus) line or at fleet terminals, for public transport.
  - Reservation and optimisation services implemented with ICT solutions.
- Electric Energy Supply for urban life:
  - Increase share of renewable generation, self-consumption (mainly heating/cooling and EV) and building optimisation.
  - Local DC-coupling of various technologies for fast charging at home.
- Regional Energy Distribution infrastructure:
  - Communication infrastructure to support self-organised local energy communities.
  - Sustainable off-grid supply with power electronics-based grid forming capabilities.
  - Virtual power plant functionality optimizing match between generation and demand.
- Operation of connected energy systems:
  - Connectivity, Security, Integrity, Resilience, Variability.
  - Interoperable platform for energy management
- Storage systems:
  - Development of grid-supporting and peak-shaving control algorithms.
  - Support for regional energy management for communities.
  - Peer-to-peer trading by using storage systems.
  - Self-powering systems for small IoT nodes.
  - Local energy harvesting to substitute battery powered devices and eliminate the high demand of energy for the battery manufacturing and distribution logistics.

3.2.3.5 Major Challenge 5: Cross-Sectional Tasks for Energy System Monitoring & Control

3.2.3.5.1 Status, vision and expected outcome

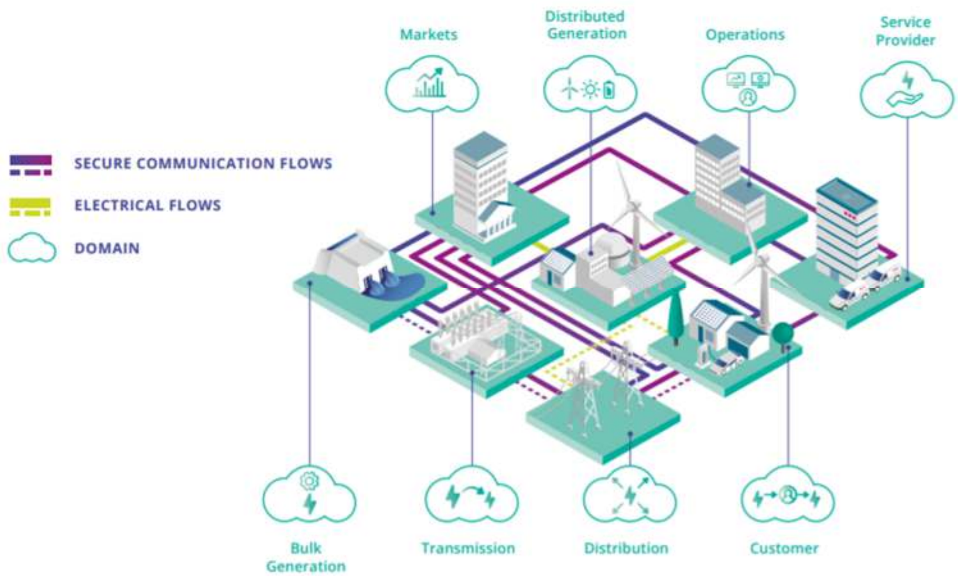


Figure 3.2.14: Interaction of actors in different Smart Grid Domains through secure communication flows and electrical flows. Source: NIST Framework and Roadmap for Smart Grid Interoperability Standards.

Focusing on current energy management platforms, they still have shortcomings in terms of automation, interaction and intelligence. Thus, when the traditional energy grid is evolving into a smart grid, it needs to integrate ICT and power electronics massively. The ECS empower the electrical utilities providers and consumers, improve efficiency and availability while constantly monitoring, controlling, and managing the demands. The huge complex networks need cross-sectional approaches for monitoring and control to achieve efficiency, security and reliability of the communication and electrical flows (Figure 3.2.14) - all based on new ECS technologies.

Optimisation in Monitoring and Control

To ensure security, reliability and stability of the total energy system, it is important to know the current state of the system at all times. Therefore, observability and state estimation together with forecast of expected production and consumption play an important role. This requires automation of the grids, use of sensors at different levels, storage of data, AI and machine learning to operate the grids in an optimised way and at the same time obeying data security and GPDR. Data collection within the grid needs to be limited on chosen parameters to avoid unnecessary costs and complexity. The IoT technology as application in the smart power grid can help to achieve sustainable energy, low latency, and reliability.[9]
Machine-learning used for forecasting energy demand in smart grid environment contributes to medium-term and long-term prediction of consumption and production and is able to solve energy management

9 Jaradat, M., Jarrah, M., Bousselham, A., Jararweh, Y., & Al-Ayyoub, M. (2015). The Internet of Energy: Smart Sensor Networks and Big Data Management for Smart Grid, *https://hdl.handle.net/10356/81241*

issues through improved accuracy of prediction[10]. It allows administrators to optimise and plan their resources and manage energy inconstancies and variations.
Nevertheless, security concerns and vulnerabilities need to be identified in today's electricity grid and sufficient solutions implemented to reduce the risks to an acceptable secure level.[11]

Energy Management Platforms for integrated energy systems

The European electrical power system is undertaking a transformation process driven by targets towards renewable energy sources. A challenge will be that all different energy infrastructures (electric, thermal, molecules) will be interconnected on high, mid and low power/voltage scale but all have completely different time scales of response. Different energy sources from different operators can be managed through ML, algorithmic trading, agile transformation, etc. In this way, challenges of current and future applications like the energy transition and the digital revolution can be faced appropriately.[12]
Energy Management Systems (EMS) are required to enable efficient and combined operation of multiple energy systems and components. Within a study that quantitatively examined 98 scientific papers dedicated to EMS in buildings and households, the identified focus areas were mostly the reduction of energy costs or peaks, as well as the increase of comfort. Results show that high computation time is a significant weakness of current EMS. A possible solution to that could be heuristic algorithms. Furthermore, the study suggests that stronger focus on high uncertainties and robustness is needed in order to transfer EMS with operational management and scheduling into practice. The integration of forecast methods also needs further attention. Regarding sector coupling (e.g. heat and electricity), major challenges exist due to great complexity and uncertainties over longer optimisation horizons. Moreover, multi-level EMS in combination with cloud computing offer exciting approaches for new research questions.[13]

Hardware

Electrical grids aim to become more distributed, smart, and flexible to meet the increasing electricity demand. For new grids, the trend is to design energy generation and consumption areas together, in distributed form. Therefore, especially power electronic devices play a crucial role to regulate distributed generation and dispersed energy-storage devices together and into the grid. Future power converters also act as edge devices actively contributing to a stable grid either in grid forming devices, virtual inertia and other functions. Hence, the intensive use of power electronic converters in the microgrid brings their control methods to the forefront, which should meet good dynamic response and high reference tracking characteristics[14]. The domain of combining low power and high-power components does require fundamentally new HW solutions. It necessitates heterogeneous integration

10 Ahmad, Chen 2018: Potential of three variant machine-learning models for forecasting district level medium-term and long-term energy demand in smart grid environment. https://www.sciencedirect.com/science/article/abs/pii/S0360544218313811

11 Aloul, Al-Ali, Al-Dalky,Al-Mardini,El-Hajj 2012: Smart grid security: Threats, Vulnerabilities and Solutions.

12 Camponesci et al. (2020). ENEL Energy Management Evolution in a growing complexity of the Italian market context. https://ieeexplore.ieee.org/document/9241132

13 Schminke (2021). Overview of the current state of research on characteristics and algorithms of energy management systems in households and buildings. https://doi.org/10.1002/er.6738

14 Bayhan, Abu-Rub (2020). Smart Energy Mangement System for Distributed Generations in AC Microgrid. Quatar Environment and Energy Research Institute, Hamad Bin Khalifa University; Texas A&M University at Qatar, Doha, Qatar.

at the highest and most diverse levels, which leads to unprecedented EMC and thermo-mechanical concerns. It may open the door to developments possible in no other application field. Exemplary, while sensors (e.g. for self-monitoring) placed directly into power switches controlling the energy flow to an entire city, two heterogeneous worlds meet (e.g. kV and pW, MA and nA). The sensors must be able to withstand strong magnetic field changes and temperature fluctuations (300 degrees +), thus requiring research and innovation.

3.2.3.5.2 Key focus areas in the cross-sectional tasks

- Self-adaptive control based on Artificial Intelligence / Machine Learning:
    - Data driven analytics (descriptive, diagnostic, predictive, and prescriptive) in smart grid.
    - Fraud detection.
    - Design, development, and application of deep learning in smart grid.
    - Artificial intelligence in advanced metering infrastructure.
    - Predictive and condition-based maintenance concepts resulting in reduced maintenance costs and increased lifetime for equipment and infrastructure.

- Algorithms for status, prediction & demand:
    - Multiobjective optimisation algorithms in smart grid; e.g. forecasting of generation and consumption.
    - state-estimation based on measurement values, simulation values, trained models (machine learning).
    - optimal utilisation of storage systems (community storage, private storage, industrial storage systems) based on technical parameters, demand and generation forecasts, customer preferences.
    - Short/long-term demand and generation forecast algorithms for different energy domains (electricity, warm water consumption, etc.) and integration into overall systems.
    - New theories and applications of machine learning algorithms in smart grid.
    - Data management, weather forecast, energy use forecast with a time horizon of 24 hours and with resolutions of at least 15 minutes (prevalent use of renewable solar, wind, hydroelectric sources according to demand profiles and use cases).

- Flexibility in management of energy supply and price offers to control the demand and avoid grid congestion
- Technologies for an open distributed energy market

- IT security, connectivity, integrity:
    - Artificial intelligence techniques for security.
    - Smart, secure edge devices for secure data management and control.

- Energy management systems for low-power/low-cost devices.
    - Smart edge computing and AI for autonomous energy control.

- Hardware Innovation:
    - H-bridge quasi-impedance source inverter (qZSI) for PV Systems.
    - Three-phase back-to-back inverter for Wind Energy Conversion Systems.
    - Ultra-capacitor with high efficiency (95%) and high-power density.
    - New generation of Smart Meter.

3.2.4 TIMELINE

MAJOR CHALLENGE	TOPIC	SHORT TERM 2025–2029	MEDIUM TERM 2030–2034	LONG TERM 2035 AND BEYOND
Major Challenge 1: Smart & Efficient - Managing Energy Generation, Conversion, and Storage Systems	Topic 1.1: smart electronic control systems for energy conversion and storage	High efficiency converters, smart actuators & sensors, Plug- and Play Functionality, Real Time Digital Twin, Integrated Security System, Status & Health Monitoring, Integrated reference communication interface, self-powering systems for off-grid operation	Further development of intelligent power devices and electronic control towards higher system energy efficiency, lower system costs and integration or newly developed device technologies - 55% GHG emissions	Getting closer to zero emissions (due in 2050)
	Topic 1.2: optimised storage possibilities	Control interfaces to batteries, fuel cells, electrolizers; Optimised converters Sensor solutions for cell and module monitoring Battery management systems Self-powered electrochemical energy storage systems (SEESs)	Grid Integration Further development based on the needs and opportunities by larger volumes	Development of excellent storage possibilities to balance energy generation volatility; efficient energy distribution and usage
	Topic 1.3 electric drives for domestic, commercial & industry application	Heat pumps, cooling devices, HVAC development, innovation and installation	Supplying clean, affordable, and secure (made in Europe) energy to these applications	"In all cases, the 2050 target is to electrify these [...] processes with technical solutions based on renewable ("clean") sources." (Green Deal)

Major Challenge 2: Energy Management from On-Site to Distribution Systems	Topic 2.1: stable and resilient multi-modal energy management systems	Distributed Generation, Interconnectivity: Renewable energy sources and grid connection	Integration of electricity, heating, cooling, and transport Virtual power plant functionality optimzing match between generation and demand; Secure gateways allowing energy trading, Coupling with energy trading systems (e.g. local energy market platforms) Renewable energy certification (labeling	Efficient energy distribution and usage; cost efficiency; high level IT-security
	Topic 2.2: energy management systems for industrial and residential customers	Development of beyond-state-of-the-art techniques for scheduling controllable loads and generators, and to forecast the weather to produce accurate generation profiles Handle uncertainties at industrial sites through ECS	optimisation module, demand and generation forecast, customer preferences, weather forecasts and price/tariff information/forecast; Demand side management for buildings Virtual Energy Market	Energy Management Systems optimizing operation of components for lifetime &revenue
	Topic 2.3: autonomous control systems	Control of high demand loads for efficient energy distribution	Price-control systems Storage devices provide flexibility, stability and reliability in the grids	Minimise costs, provide peak-shaving; hybrid solutions; novel grid architectures for manufacturing to enable adaptive production optimisation

Major Challenge 3: future transmission grid	Topic 3.1: Grid stability during the industrial transition	Development of components for HV transmission for 1-2 MV or even higher voltages New solutions for high power electronics, combined with sensors and ICT for monitoring, control and protection	Further improvements on grid capacity with highest possible efficiency Development of business models that encourage new technological solutions	European Energy transition to zero-carbon emissions
	Topic 3.2: Resilient systems for the European transmission grids	Water-resistant components/modules Autonomous electricity generators based on fuels cells Modelling Intelligent power devices, systems, and switches	Modelling, weather forecast, sensing data Digital twin ECS for multi-modal energy systems (cooling with excessive energy, use thermal capacities)	Multi-Modality across Europe
Major Challenge 4: Achieving Clean, Efficient & Resilient Urban/ Regional Energy Supply	Topic 4.1: Regional energy distribution infrastructure	Secure Cross Regional Transmission Infrastructure communication infrastructure to support self-organised local energy communities	Sustainable off-grid supply with power electronics based grid forming capabilities	Energy flows between sectors and their storages ensure the highest use of renewable energy while balancing fluctuations
	Topic 4.2: Electric energy supply for urban life and mobility	Development of household and public charging infrastructure; charging points on bus lines or terminals Reservation and optimisation services implemented with ICT solutions. Bi-directional charging and grid stabilization solutions	Increase share of renewable generation, self-consumption (mainly heating/cooling and EV) and building optimisation Local DC-coupling of various technologies for fast charging at home Sector coupling of large energy users (e.g. harbours, airfields) with urban users (electric cars, households,	Emission free cities with electrification and decentralised storages to improve efficiency and reliability

	Topic 4.3 Storage systems for urban communities	Development of grid-supporting/forming and peak-shaving control algorithms Battery energy/ heat/ cooling storage devices for households Integration of fuel cells and electrolyzers	Peer-to-peer trading by using storage systems; Self-consumption optimisation	Support for regional energy management for communities

Major Challenge 5: Cross-sectional Tasks for Energy System Monitoring & Control	Topic 5.1: AI, machine learning and algorithms for status, prediction and demand	Data driven analytics and deep learning in smart grid; AI in advanced metering structure; smart sensors with improved data processing; stream processing for real time application	Innovative approaches ensuring clean, secure and affordable energy for EU citizens; multiobjective optimisation algorithms in smart grid; optimal utilisation of storage systems; short-/long-term demand and generation forecast algorithms for different energy domains	Safe and interconnected smart grid netword; cross-sectional approaches for energy monitoring and control; integrated energy systems; optimal match between generation and demand; energy flexibility
	Topic 5.2: IT security, connectivity, integrity	Smart, secure edge devices for secure data management and control	Artificial intelligence techniques for security	Eliminate security vulnerabilities as best as possible
	Topic 5.3: Hardware	Improvements in robustness of HW devices to withstand strong magnetic field changes and temperature fluctuations	Good dynamic response and high reference tracking characteristics of power electronic converters; new HW solutions to combine low power and high power components	Optimal regulation of distributed generation and dispersed energy-storage devices; robust devices able to control high energy flows

3.3

ECS Key Application Areas

DIGITAL INDUSTRY

3.3 Digital Industry

"Digital Industry" chapter intends to highlight the evolution of traditional industry through the introduction of cutting-edge digital technologies in production, management and distribution processes. These technologies include the Internet of Things (IoT), artificial intelligence (AI), data analysis, co-robotics, 3D printing and all digital innovations that can be usefully adopted in industry (advancements on sensing, powering, communications, computing, etc…)

Digital Industry is characterized by the increasing integration of the physical world with the digital one by means of the data exploitation provided by an ever more pervasive digitalization. The digitalization allows to move under the same objective heterogeneous topics inherent to the different steps on the lifecycle of the products -like as from the design to the efficiency and the maintenance of the production machineries- involving all aspects of industry and business activities, including workers and the working environment.

3.3.1 Scope

To be able to manage everything in a machine, factory or company network, industries have divided the necessary technologies into levels or technology stacks. In these levels or stacks, sensors and actuators are closest to processing materials or handling items, and therefore seen as the lowest in the hierarchy in the Edge-to-Cloud Continuum. Moving up the levels, you find super-sets and/or System of Systems and/or interconnected IoTs like devices, unit processes, production or manufacturing lines, operations control, company or enterprise business processes, and an increasing number of machines, lines, company boarders, as manufacturing has become more networked, complex, dynamic and global.

The Industry 4.0 changes to the mode of operation have a profound impact on how are managed and operated the factories, supply chains, construction zones and processes. Powerful net-worked digital tools are needed to achieve the necessary Situational Awareness and control of autonomous vehicles, robots and processes at various autonomy levels allowing the ever more pervasive Industry 5.0 approach. The technological tools that are part of the symbiotic decentral-ized Web 5.0 concept can encompass such complex non-hierarchical environments and allows control of user data.

The European Industry 4.0 and 5.0 revolution needs an open-source-based, stable and extensible semantic Internet to overcome user interface, networking and communication challenges. The Internet with smart agents is an overarching concept containing multiple technologies, such as Industrial Internet of Things, Artificial Intelligence (AI), Advanced Data Analytics, Augmented and Virtual Reality (AR /VR). The Internet with Industrial Metaverse (Digital Twins) can be seen as an expanding set of interconnected virtual and augmented worlds, accessible from AR/VR head mounted displays, desktop computers and even mobile phones. The entities that are present in the Industrial Metaverse may be humans, devices and autonomous physical and virtual machines such as Digital Twins and AI agents. The digital transformation of the manufacturing processes, creation of the software tools, data-driven and AI-supported decision making, harmonization of

human-technology relationship, is enabling upcoming human-robot coworking toward Industry 5.0.
This SRIA addresses the digitalisation of the major European industrial sectors promoting the European sovereignty in the internal manufacturing ecosystem together with future sustainability and greener industrial processes and artefacts. These include discrete manufacturing (e.g. manufacturing of automobiles, trains, airplanes, satellites, white goods, furniture, toys and smartphones), process industries (e.g. chemical, petrochemical, food, pharmaceuticals, pulp and paper, and steel), provisioning, production services, machinery and connected machines, UAVs and robots. Emphasis is also given to any type of factories, productive plants and operating sites, value chains, supply chains and lifecycles, new materials for structures and electronic components.

Digitalisation is as a key enabler for the future success of European industry. This Chapter will address the potential for the development of topics such as responsive, smart and sustainable production, Artificial Intelligence (AI) in digital industry, industrial services, digital twins and autonomous systems. As discussed at the end of the Chapter, nearly all of the topics in the Technology chapters of the SRIA are of vital importance to industrial applications. These include standardisation, engineering tools, cybersecurity and digital platforms. To digitise European industry and enlarging and embedding the related e-supply chains, potentially all enabling technologies will need to be employed to realise the required competitive edge, and of course a focus on digital industry would not be complete without the exploitation of enabling technologies in the industry field.

Today, the digital landscape remains fractured, with significant challenges in areas such as standardisation, interoperability, and translating research to real commercial impact. These challenges must be met effectively if we are to achieve a strong, greener, resilient, responsive European economy, where sustainable, human-centric solutions help Europe achieve strategic autonomy into the future, as also well outlined in the Draghi Report[1] “EU competitiveness: Looking ahead” and in the more sectorial WMF Report 2024 called “New Perspectives for the Future of Manufacturing: Outlook 2030”[2].

3.3.2 Application trends and societal benefits

The recent “2030 Digital Decade” Annex I[3] of the Shaping Europe’s Digital Future[4] Commission Report highlights the urgent need for the EU to prioritise action in areas that foster innovation and growth, improve productivity, and mitigate disruptions – in particular in the areas of digital technologies and digital skills.

---

[1] https://commission.europa.eu/topics/strengthening-european-competitiveness/eu-competitiveness-looking-ahead_en

[2] https://www.effra.eu/news/wmf-report-2024-new-perspectives-for-the-future-of-manufacturing-outlook-2030

[3] Annex_1_Competitiveness_and_sovereignty_people_smart_greening_policy_coherence_and_synergies_dKR4myZOv5WrJuPEHds4ip3XXOA_106688.pdf

[4] https://digital-strategy.ec.europa.eu/en/policies/2024-state-digital-decade-package

Europe needs an industry that becomes greener, circular and more digital while remaining competitive on the global stage. The twin ecological and digital transitions will affect every part of our economy, society and industry. They will require new technologies, with investment and innovation to match. They will create new products, services, markets and business models. They will shape new types of jobs that do not yet exist which need skills that we do not yet have. And they will entail a shift from linear production to a circular economy.

The digital sector will also contribute to the European Green Deal, both as a source of clean technology solutions and by reducing its own carbon footprint. Scalability is key in a digitalised economy, so strengthening the digital single market will underpin Europe's transition. Europe must also speed up investment in research and the deployment of technology in areas such as AI, 5G/6G, data and metadata management. 5G/6G private networks will become into the industry and at the same time hybrid clouds will enable networks that are sharing content and resources seamlessly. This will be required by Web 4.0 smart services and things: goal oriented intelligent connection of semantic resources. Connectivity will expand between systems and people. Personnel at the mill level will be more aware and artificial knowledge from the cloud will be nearby all the devices, sensors and processes to the human. This can be made more accurate, for example, with precision time protocol (PTP) support on hardware level (industrial devices at least) to get better indoor location (by means communication improved capabilities, like as Bluetooth Low Energy (BLE), Ultra Wide Band (UWB), and advancements in sensors, like as time of flight).

European factories and machines already have a high level of automation and digitisation. Many of the leading end-user companies are European based, and Europe also has a number of significant system and machine building, engineering and contracting companies that have a competitive edge in automation and digitisation. The business environment is changing. Through specialisation in new or niche end products, production is becoming more demand-driven and agile, while production is increasingly geographically distributed enlarging its supply chain. In addition, the outsourcing of auxiliary business functions such as condition monitoring and maintenance is gaining in popularity, leading to highly networked businesses. There are many opportunities for energy, waste, water, material, recycling optimisation, etc., over the value chains and across company boundaries. Such advantages can only be realised by having a significant increase in digitisation and embedding functions and intelligence within the production physical layer.

This transition should include the adoption of applications that do not require to be kept internal or confidential, as solutions based on web/cloud services allows for the mediation of key factors such as their use by non-AI professionals, and off-line development of advanced criteria models and inferential engines through the expertise of specialised centres.

The exploitation of AI for core business functions generally requires a complete rethinking of data management and their use and tracking inside the supply chain. Instead, the implementation of a System of System (SoS) framework enables the data to be to capitalised on through appropriate actions, in which analysis and analytical tools usually reach their limits. Interaction between systems will be more direct (knowing neural nets). This will require more standardized interfaces that will enable adaptive connectivity with secured access rights (part of information will remain always hidden).

Industrial Metaverse, based on digital twins, will speed up training of employees, and furthermore, facilitate building required trust to use intelligent agents. It will enable virtual simulations before actual building physical system and speed up research and innovation cycles It can be used to help field workers with AR/VR. Even product design collaboration around the world will come possible. Virtual designs will enable building physical products and vice versa: adding digital assets from physical items.

The actual value chain will come from existing installations, as it is unusual for new factories to be built. As new, fast and secure communication protocols will provide easy connectivity and interoperability across systems; this will enable the potential for extensive integration. Easy access to a secure internal network will provide all existing information to users at anytime and anywhere within the plant. Moreover, new interesting features could be accessed through cloud or edge-based computing systems. However, this will require new hardware infrastructure to be added to the plant, along with greater processing power to handle larger amounts of data.

Digital infrastructure and micro services will help evolve business models towards selling added value as a service. Investment in projects will create networks between vendors and providers. In modern business- to-business (B2B) relationships, ongoing R&D and industrial pilots will aim to deliver a range of after-sales services to end customers. Typically, such services will include condition monitoring, operations support, spare parts and maintenance services, help desks, troubleshooting and operator guidance, performance reporting, as well as the increasingly required advanced big data analytics, prognostics-based decision support, and management information systems.

Industrial services often represent 50% or more of industrial business volume, and this share is steadily growing. The share of services is generally higher in high-income countries than in low-income countries. The importance of service businesses in the future is obvious, since they also enable sustained revenue after the traditional product sales, with the service business being typically many times more profitable than the actual product sale itself.

3.3.3 Major Challenges

Each Major Challenge (MC) has a focus on specific topic, but we have to highlight how all digital industry MCs are characterized by a high grade of interaction among them, in which each MC takes advantages on improvements of topic in another MC. Artificial intelligence, robotics, digital twins, augmented reality, etc., appear together as connected tools to improve industry in every possible part of its domain.

On Digital Industry domain, six Major Challenges have been identified:

- **Major Challenge 1**: Responsive and smart production. Focusing the Efficiency
- **Major Challenge 2**: Sustainable production. Focusing the Planet
- **Major Challenge 3**: Artificial Intelligence in digital industry. Focusing the Smartness
- **Major Challenge 4**: Industrial service business, lifecycles, remote operations and teleoperation. Focusing the Resilience of Supply Chains
- **Major Challenge 5**: Digital twins, mixed or augmented reality, telepresence. Focusing the Digitalization
- **Major Challenge 6**: Autonomous systems, collaborative robotics. Focusing the Autonomous Production

3.3.3.1 Major Challenge 1: *Responsive and smart production*

3.3.3.1.1 Status, vision and selected outcome

Responsiveness, flexibility and smartness is currently considered inside the European productive landscape based on pre-Industry 4.0 automation technology. At the same time, in addition to the critical situations to be handled, European Industry must be much more responsive to any changes that may occur, not only from a scientific and technological point of view.

Today's automation solutions are not inherently flexible by design.
Adaptive and self-learning systems principles must be applied to automation solutions, enabling the automation system and solution flexibility, responsiveness and smartness.

Many European initiatives and reports cover this topic:

- "Align Act Accelerate" Report[5] prepared by European Factories of the Future Research Association (EFFRA[6]) Expert Panel: a proposal of comprehensive reform of the Eu Research & Innovation Funding.
- SMART-EUREKA[7] mission-vision to boost the competitiveness, growth and attractiveness of the European discrete manufacturing industries through the promotion of R&D&I in an open community of industrial organisations.
- ManuFuture[8] claims that "the European manufacturing system in 2030 must be resilient and adaptive to cope with a rapidly changing and unpredictable environment, overcome disruptions and adapt to meet the changing market needs". In their 2020 report[9], the World Manufacturing Foundation calls to "leverage on AI to detect and respond to disruptions in supply networks".

The main benefits of a more responsive and resilient production are:

- Ability to forecast the evolution of the demand-offer-competition ecosystems
- Capability to rapidly change production and provisioning.
- Efficiency to become profitable with high-mix, low-volume production.
- Capacity to operate, even with decreased operational capability.
- Enable rapid concrete innovation, not exclusively via rapid integration and product evolution but also re-engineering the overall process and components production.
- Limit deterioration in performance, reliability, maintainability and interoperability when plants face disturbances.

---

[5] https://www.effra.eu/news/expert-panel-proposes-comprehensive-reform-of-eu-research-and-innovation-funding-align-act-accelerate-report/

[6] https://www.effra.eu

[7] https://www.smarteureka.com/about-us/mission-vision/

[8] http://www.manufuture.org/wp-content/uploads/Manufuture-Vision-2030_DIGITAL.pdf

[9] 2020 World Manufacturing Report - World Manufacturing Foundation

Chips JU will play a lead role in this evolution to more responsive, robust and resilient factories. From sensors integrated in wearables and prosthetics, to SOS that enable self-healing and self-reconfiguration, responsive and resilient manufacturing has always been an important challenge.

AI will also play a key role in increasing the flexibility of manufacturing systems. One example is AI applied to real-time scheduling that allows a production process to flex around rush orders and disturbances in the line or supply chain. However, while the current trend of deep learning is opening up limitless possibilities in some areas, there is a need to apply other AI approaches that are more explainable, more aware of the environment and the task at hand.
The still suffering tail effect of Covid-19 and the running wars in Ukraine and Gaza with the worldwide impacts expected for the future years has highlighted many of the reasons why a flexible factory needs to adapt better in times of change to be a more useful part of the European response to such crises by:

- Modifying production based on medical needs and exogenous inputs.
- Scheduling production with less human resources and social distancing constraints.
- Empowering agile working and telepresence.
- Adapting to changes in the supply chain, promoting European independence.
- Developing capacity for in-kind, or inside a shorter supply chain for the production of components that usually come from the worldwide market-based supply chain.
- Developing capacity for redesign and re-engineering, due to lack of raw materials and electronic components.

Although automation and digitisation are the building blocks for building a flexible, resilient manufacturing industry, the importance of a well-trained and agile workforce cannot be underestimated. Workforce agility and flexibility, achieved, for instance, through cross-skilling, empowered by smart technologies like AR/VR or assisting chatbots, make humans indispensable in any production process.
In terms of standardisation, standards are a significant and necessary part of all industrial applications. The modern digitalisation of industry could not exist without standards, as without standards interoperability would not be possible. They enable extensive industrial projects while ensuring quality, safety and reliability. Many engineering methods are standardised, and provide textbook consistency across professional engineering. However, standards must also be supported by the relevant engineering tools, etc., as those required for design or development are different from those required at the operation stage.

3.3.3.1.2 Key focus areas

- **Robust optimal production, scalable first-time-right production**: future manufacturing plants should become more robust in the sense they can continue production even when facing a disturbance. This will require advances in, for example, self-healing and redundant automation systems, first-time-right, zero-defect manufacturing, and predictive maintenance empowering very adaptable production.
- **Mass customisation and personalised manufacturing, customer-driven manufacturing**: progress in recent years towards lot-size-one manufacturing and personalised product design will continue to grow in the next few years: IoT sensors will self-correct for disturbances, real-time warehouse connectivity will allow to optimise the shop-floor, edge devices will perform predictive analytics. All of these advances will allow greater flexibility.
- **Resilient and adaptive production, including the shortening of supply chains and modular and flexible factories**: Resilience is a critical property for systems that can absorb internal and/or external stresses and adjust their functional organisation and performance to maintain even with a reduced productivity the necessary operations. A resilient factory will continue to operate without any breaks to achieve its objectives under varying conditions and with the ability to overcome that stress to return to ordinary condition of functionality and productivity.
- **Cognitive production**: This involves deploying both natural and artificial cognition alongside IoT devices and data analytics to enable new analytics and learning that can enable responsive and sustainable adaptable production. For example, real-time monitoring against lifecycle assessment (LCA) criteria can be facilitated by the implementation of AI. More generally, it is important for cognitive production to support the emergence of simplicity rather than the combinatorial growth of complexity when complex cyber-systems are combined with complex physical systems.
- **Manufacturing as a service**: Technological advances have the potential to expand the geographical distribution of manufacturing and facilitate manufacturing as a service, MaaS (a well- known example is 3D printing). The trend to move part of labour-intensive production into high value manufacturing exploits MaaS, allowing to outsource parts of the production chain, replacing the need to have dedicated lines and even the whole factory (factory less goods). Another interesting opportunity for MaaS is moveable factories, which circumvent the need for new industrial infrastructure. The scope for moveable factories is enhanced by the range of manufacturing machines and power sources that are becoming increasingly small and light enough to fit into trucks, trailers, carry cases, etc. MaaS empower the dynamic aspects of outsourcing, provisioning, making the supply chain ecosystem more resilient.
- **Embedded/Edge/Cloud architectures**: Nowadays system architectures mainly consist of three layers of computing devices. *Embedded computing* reside very close or attached to the machinery or process. *Near Computing* devices are often called *edge computers*, routers, or local servers. Near computer nodes are powerful computers themselves and communicate both to the embedded computing cards and to the *cloud* via internet. A

special boost in the picture comes from the **5G** technology. There, the 5G communication technology takes care of the edge-to-cloud and edge-to-edge communication which will be faster than anything before. Workload and services are moved from the centralised data centres (core network) ever more to the proximity, benefiting mainly latency which is crucial for time critical application and services.

- **Standardisation**: Due to the ongoing legacies of the many existing standards and their installed base (number of units of a product or service that are actually in use), focus should be on bridging the systems of the various standards. This should involve developing semantic technologies to master these diverse and numerous standards, including software or platforms that enable effective connectivity at a high application level, as well as respective digital testing, development environments and licencing. This is key to ensure there is wide acceptance and support of software vendors, engineering offices and end-users.

3.3.3.2 Major Challenge 2: Sustainable production

3.3.3.2.1 Status, vision and expected outcome

This **Major challenge** focuses on how Industry 4.0 should address the future regulation or market requirements emerging from the European Green Deal and zero-carbon (or below carbon-neutral) operations.

Nearly 200 countries have committed to the Paris Agreement on climate change to limit global warming to below 2ºC. Any update about new challenges and activities are regularly published by the EU Parliament[10]. The rapid transformation of all sectors is therefore required. In fact, many European countries have set even more ambitious targets, and Chips JU could have a great bearing in reducing environmental impact through sustainable manufacturing, including energy- and resource-efficiency and by applying circular economy strategies (eco-design, repair, re-use, refurbishment, remanufacture, recycle, waste prevention, waste recycling, etc).

There are some so-called rare earth metals to save, and any kind of careless materials usage can be proven uneconomical and risk to the environment. The vision of Sustainable Process Industry through Resource and Energy Efficiency (SPIRE) categorises the high-level goals discussed above into more practical action, as follows.

- Use energy and resources more efficiently within the existing installed base of industrial processes. Reduce or prevent waste.
- Re-use waste streams and energy within and between different sectors, including recovery, recycling and the re-use of post-consumer waste.

[10] https://www.europarl.europa.eu/topics/en/article/20200618STO81513/green-deal-key-to-a-climate-neutral-and-sustainable-eu#:~:text=Parliament%20adopted%20the%20EU%20Climate,global%20fight%20against%20climate%20change.

- Replace current feedstock (raw material to supply or fuel a machine or industrial process) by integrating novel and renewable feedstock (such as bio-based) to reduce fossil, feedstock and mineral raw material dependency while reducing the CO2 footprint of processes or increasing the efficiency of primary feedstock. Replace current inefficient processes for more energy consumption reduction.
- Resource-efficient processes when sustainability analysis confirms the benefits.
- Reinvent materials and products to achieve a significantly increased impact on resource and energy efficiency over the value chain.

ManuFUTURE Vision 2030 combines these objectives as shown in Figure 3.3.1

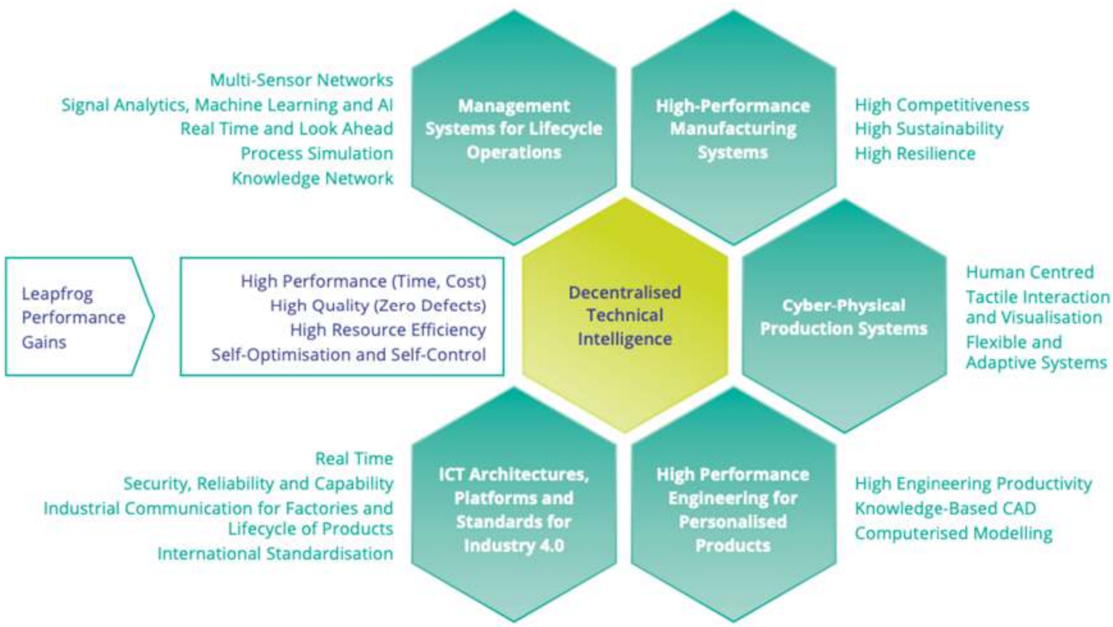


Figure 3.3.1 - The visionary manufacturing system for adding value over the lifecycle with decentralised technical intelligence (Source: ManuFUTURE, "Strategic Research And Innovation Agenda (SRIA) 2030"[11])

A new upcoming EU co-programmed partnership called Innovative Advanced Materials for Europe (IAM4EU) will be supported by the IAM-I[12] (the Innovative Advanced Material Initiative) association, which already released a template for collecting inputs for its SRIA[13]. This materials value chain initiative is of paramount importance for the EU manufacturing Industry and will have a great impact also in Digitalization and Semiconductor Design & Production.

3.3.3.2.2 Key focus areas

---

[11] http://www.manufuture.org/wp-content/uploads/ManuFUTURE_SRIA_2030_Vfinal.pdf

[12] https://www.iam-i.eu/

[13] https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.iam-i.eu%2Fwp-content%2Fuploads%2F2024%2F07%2FIAM4EU-Industrial-RI-Needs_template-2.docx&wdOrigin=BROWSELINK

- **Monitoring flows of energy, materials, waste and Lifecycle assessment**: It is already commonplace in many industry sectors (food, medicine, etc) that material and energy streams need to be fully traced back to their starting point. As more and more products, raw materials, etc, become critical, this implementation strategy must be expanded. Flows need to be monitored. Sustainable manufacturing needs comprehensive environmental data and other measurements that may have been in place when the relevant manufacturing or production was initiated. On the other hand, this is a very typical application for many types of IoT sensor and systems that can be informed by careful "Life Cycle Assessment" (LCA). LCA is a prerequisite for holistic environmental evaluation, and it is a simple but systematic method, that requires a mixed combination of extensive and comprehensive models and data.
- **Virtual AI assistants**: Discharges or losses mostly happen when production does not occur as planned, due to mistakes, the bad condition of machinery, unskilled operation, and so on. Human factors cause most of the variation in the running of continuous processes. There should therefore be a focus on how to implement an AI-supported assistant to help operators by providing advice and preventing less than optimal changes and addressing the dynamicity of the production environment.
- **Human–machine interfaces and machine-to-machine communications**: Augmented reality (or virtual reality) will be used to support a number of tasks. Enhanced visualisation of data and analytic results will be required to support decision-making.
- **Human operators in more autonomous plants and in remote operations**: the relationship between machines and the human factor needs to be rethought. In terms of the logic of human-machine interface, from touch displays, to wearable devices and augmented reality, but also to maintain the centrality of the human factor within the new contexts. The "Skills 4.0" are necessary for the management of new technologies for data administration, for privacy, for cybersecurity and much more.
- **Human safety**: With the localisation of personnel, machines and vehicles, situation-aware safety (sensing of safety issues, proximity detection, online human risk evaluation, map generation, etc) will become increasingly vital.
- **Competence and quality of work in a human-centred manufacturing**: At a strategic level, the European automation and industrial IT industry depends on its ability to attract skilled personnel to maintain their competence over time. A higher level of formal training may be required for workers in production and maintenance. Greater specialisation is constantly introducing products and processes that require greater company-specific training.
- **Green Deal**: Policy initiatives aimed at putting Europe on track to reach net-zero global warming emissions by 2050 are key to the European Commission's European Green Deal. Following the highly challenging objectives of the Green Deal, all industries must focus on high efficiency, low energy usage, carbon-neutrality or zero-carbon usage, zero waste from water, soil and air – all measured, calculated or estimated on product, factory, global and lifecycle levels. European industry must research and discover new materials while paying a great deal of attention to recycling, re-use, and de-manufacturing and re-manufacturing. As a "niche" and not exhaustive example new RICS-V based computing

hardware will be needed to reduce energy used at data centres. Extra 3 x performance will be gained compared to ARM Cortex-A75.

Many of these advances will require extensive development in the other engineering, business or social domains, even at the individual level, that are outside of the Chips JU focus. However, it is also obviously the case that a growing part of these approaches will be implemented through the significant help of electronics and software technologies. The need for Chips JU technologies is diverse, and it is not useful to indicate one single technology here. High performance, high precision, careful and professional engineering and decision-making are needed – often at a much higher level than today.

3.3.3.3 Major Challenge 3: Artificial Intelligence in digital industry

3.3.3.3.1 Status, vision and expected outcome

Major challenge 3 focuses on connected and smarter cyber-physical systems (CPS), industrial internet, big data, machine learning, optimization, and other AI-based methods. Local edge-based intelligence is seen as an opportunity for Europe. AI optimized and Open hardware becoming more important to support European AI Framework. This Major challenge extends toward AI-enabled, adaptable, resilient factories, including the human as a part of a "socio-technical" system. AI methods, extending various systems used in modern industries, like(predictive) condition monitoring and maintenance, scheduling, planning, or quality control, will be applied to not only support reconfigurable first-time-right/zero-defect manufacturing, but also to support human decision-making (considering uncertainties). Consequently, AI-enhanced systems will enable resilient manufacturing ecosystems based on new business models, increase safety & security in working environments, and improve productiveness and quicker return from investments. In this context, Explainable AI (XAI) is another an emerging field for better understanding of decision making to increase human trust and robustness of emerging technologies. An important challenge here is to lead not only the digital transformation of Industry 4.0/5.0, but also the next generation of Chips JU platforms supporting AI-driven human-centric autonomous Industry 4.0/5.0 operations. Condition monitoring techniques can be applied to many types of industrial components and systems, although often at additional cost. Commonly, the business value required from condition monitoring depends on the higher availability of equipment and, for production processes, information provision to be able to plan and act on maintenance proactively instead of reactively, as well as to offer decreased cost and improved on-time delivery. Other business values that may be of interest are safety and the optimal dimensioning/distribution of spare parts and maintenance staff. Thus, serious breakdowns and unplanned interruptions to production processes can largely be avoided using condition monitoring.

AI will impact several main areas, all of which are relevant to Digital Industry improving:

- Productivity, by exploiting AI in the design, manufacturing, production and deployment processes.
- Flexibility, by using AI throughout the value stream, from supply to delivery, increasing the autonomy and resilience of each process in the value chain.
- Customer experience, by using AI to make products faster and with better quality, and to provide more efficient services.
- Assistance to human operators in circumstances of rising complexity by using AI to support the decision-making process with ever-increasing levels of complexity and dynamics.
- AI radical innovations would enable more productivity and new products that just supporting, maintaining and improving existing ones

These fundamental impacts will be experienced in all areas of every market sector. In manufacturing and production, AI will deliver productivity gains through a more efficient resources, energy and material use, better design and manufacturing processes, and inside products and services, enhancing their operation with more refined contextual knowledge.

The agenda here is cross-sectorial, focusing on AI applied in any domain. However, the impact of AI in Digital Industry is of particular significance. AI in manufacturing is for example in one of the main focus of AI, Data and Robotics[14] (ADR)and area of expected investments, already roadmapped by ADR since 2021[15]

3.3.3.3.2 Key focus areas

a) European AI framework

Figure 3.3.2 sets out the context for the operation of AI public/private partnerships (PPPs), as well as other PPPs or joint undertakings (JUs). It clusters the primary areas of importance for AI research, innovation and deployment into three overarching areas of interest. The European AI framework represents the legal and societal fabric that underpins the impact of AI on stakeholders and users of the products and services that businesses will provide. The AI innovation ecosystem enablers represent essential ingredients for effective innovation and deployment to take place. Finally, the cross-sectorial AI technology enablers represent the core technical competencies that are essential for the development of AI systems.

[14] https://adra-e.eu/events/dfki-organizes-1st-adra-e-cross-project-workshop

[15] Strategic Research, Innovation and Deployment Agenda, AI, Data and Robotics Partnership, September 2020

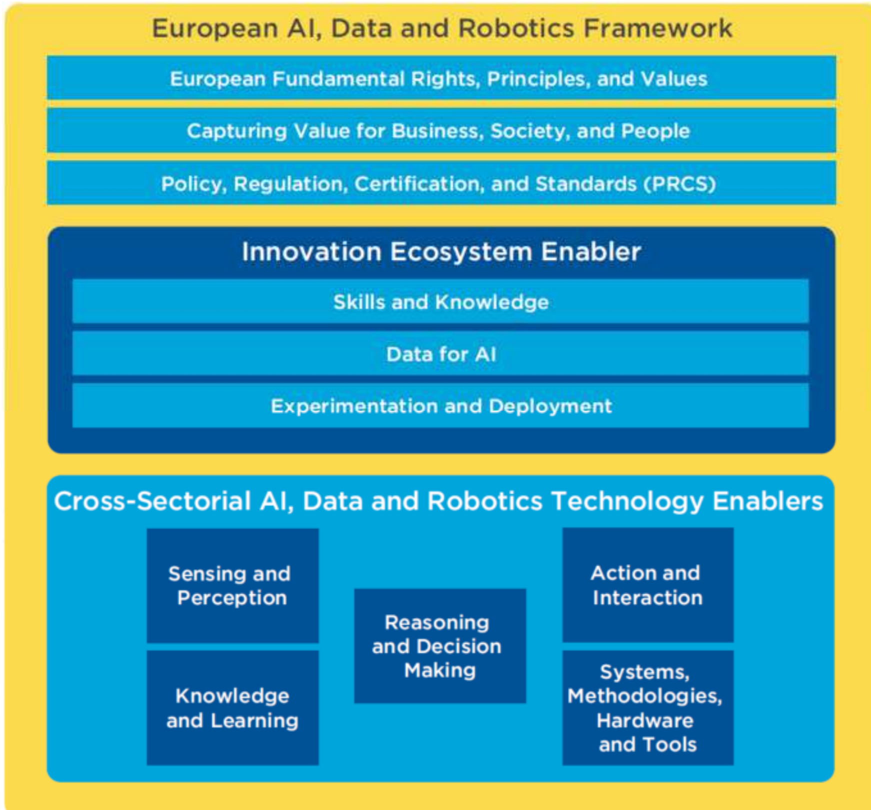


Figure 3.3.2 European AI framework and enablers – SRIDA on AI, Data and Robotics Partnership[16]

In June 2024 a big step forward has been made by the EU Parliament publishing the AI Act[17], which is the first European regulation on Artificial Intelligence. Nevertheless ADRA (the AI, data and Robotics association) published in 2023 the "Strategic Orientation towards an AI, Data and Robotics Roadmap 2025-2027"[18]. Hence the European AI framework is still moving forward.

b) AI in manufacturing

- AI for dynamic production planning and management: This involves taking real-time decisions to optimise the factory operation by quickly modifying the productions schedule, based on the current state of the shop floor, predicted sales orders, unexpected events such as machine breakdowns or changes in job priorities, etc.
- Virtual models spanning all levels of the factory life and its lifecycle: A holistic and coherent virtual model of the factory and its production machinery will result from the contribution and integration of modelling, simulation and forecasting methods and tools that can strategically support manufacturing-related activities.

[16] https://bdva.eu/download/92/publications/3122/ai-data-and-robotics-sirda-joint-version-rel3-1.pdf

[17] EU AI Act: first regulation on artificial intelligence | Topics | European Parliament (europa.eu)

[18] https://www.adr-association.eu/sites/default/files/2024-04/ADRA-roadmap-May2023_PostConsultationVersion.pdf

- AI for green/sustainable manufacturing**:** The development of software-based decision-support systems, as well as energy and resource management, monitoring and planning systems, will lead to overall reduced energy consumption, more efficient utilisation and optimised energy sourcing.
- AI applied in supply chain management**:** Planning and managing logistics for real-time operations, collaborative demand and supply planning, traceability, and execution, global state detection, time-to-event transformation, and discrete/continuous query processing would therefore be a challenge in view of the distributed nature of these elements.
- AI for advanced manufacturing processes**:** The ability to design functionality through surface modifications, functional texturing and coatings, enabling improved performance, embedded sensing, adaptive control, self-healing, antibacterial, self-cleaning, ultra-low friction or self- assemblies, for example, using physical (additive manufacturing, laser or other jet technologies, 3D printing, micromachining or photon- based technologies, physical vapour deposition, PVD) or chemical approaches (chemical vapour deposition CVD, sol–gel processes) will deliver high functionality and hence high-value products.
- AI for adaptive and smart manufacturing devices, components and machines**:** Embedded cognitive functions for supporting the use of machinery and robot systems in changing shop floor environments. Open Hardware based AI solutions will also support European AI Framework

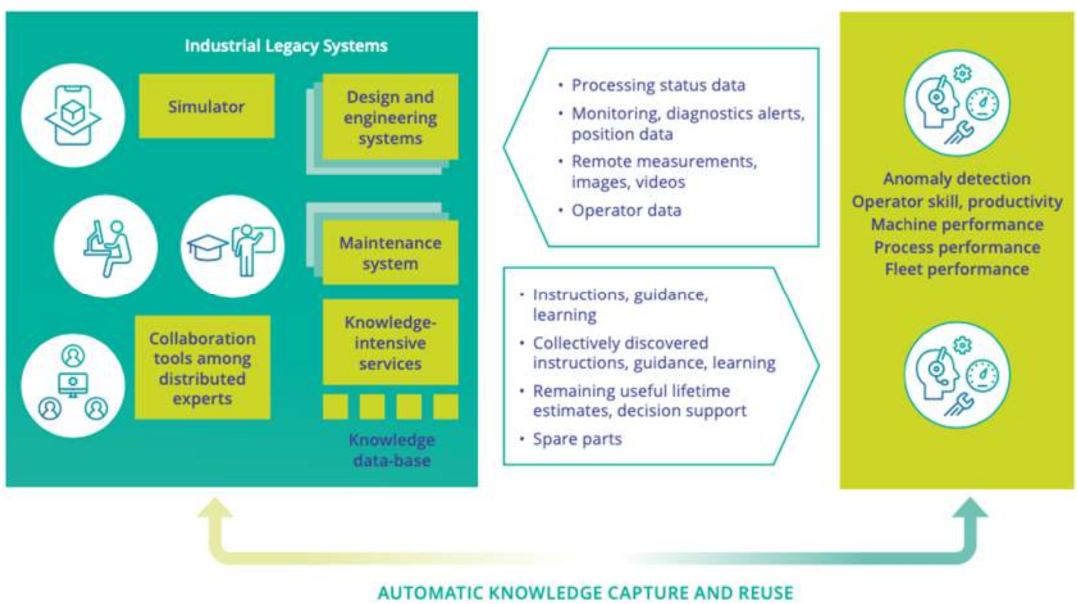


Figure 3.3.3 - Industrial service business between a machine or system vendor or service provider and an end-customer. Services or lifecycle businesses deal with, for example, anomaly detection or condition management, operator skills development, productivity issues, machine or system performances, and fleet performances.

c) AI for decision-making

Decision-making is at the heart of AI. Furthermore, Explainable AI will increase "human-centric" approaches in autonomous industry.

- AI can support complex decision-making processes or help develop hybrid decision-making technologies for semi-autonomous systems.
- Human decision-making, machine decision-making, mixed decision-making and decision support.
- Sliding or variable decision-making, dealing with uncertainty.
- AI for human interaction with machines.

d) AI for monitoring and control

- AI for control technologies.
- AI for monitoring services.
- AI for maintenance systems for increased reliability of production systems.
- AI services for continuous evaluation and mitigation of manufacturing risks.
- AI for quality inspection.

3.3.3.4 Major Challenge 4: Industrial service business, lifecycles, remote operations and teleoperation

3.3.3.4.1 Status, vision and expected outcome

The volume and value of industrial services are increasing by between 5% and 10% every year. The share of services has exceeded the share of machinery for many machines, system and service vendors – not just for a final assembly factory, but also for companies in supply chains. Companies are willing to take larger shares of their customers' businesses, initially as spare part suppliers, but increasingly for remote condition monitoring, as well as extending this to a number of those tasks previously considered as customer core businesses. From a customer point of view, such a shift in business models lies in the area of outsourcing.

Industry as services is changing the production through use of externalisation moving local tasks to external and ever more specialised providers, benefiting of greater flexibility, resilience and adaptability on production; a new proposition of supply chain is embedding all production phases, from the procurement of raw materials and semi-finished products up to the customer services and/or the design of parts or whole final product.

While many businesses have become global, some services are still provided locally, at close to customer's locations, while other services are provided centrally by the original vendor or companies specialised in such services. Similarly, as there may be extensive supply chains underpinning the vendor companies, the respective services may also extend to supply chain companies. The industrial era is becoming a service era, enabled by high-end Chips JU technologies. This distributed setting conveniently fits into modern edge-to-cloud continuum innovative architectures as computing power engines and infrastructures enabling emulation, training, machine learning and communication platforms enabling real-time interconnections.

The importance of service businesses to the future is evident as they enable a revenue flow beyond traditional product sales, and more importantly they are typically much more profitable than the product sales itself.

The service business markets are becoming more and more challenging, while high income countries are focusing on the high-skilled pre-production and lifecycle stages. Fortunately, in the global service business market, Europe can differentiate by using its strengths: a highly skilled workforce, deep technology knowledge and proven information and communications technology (ICT) capabilities. However, to ensure success it needs new innovations and industry-level changes.

3.3.3.4.2 Key focus areas

- **Remote operations, teleoperation**
    - Remote engineering and operations, telepresence: Operating or assisting in operations of industrial systems from remote sites.
    - Edge/cloud solutions: Implementing distributed service applications on effective edge cloud systems.
    - 5G with very low latency will be used for remote operations.
- **AI Services for monitoring and collaboration**
    - Collaborative product-service engineering, lifecycle engineering: Extending R&D to take into account how products and systems will be integrated into the industrial service programme of the company. This should possibly be enhanced by obtaining further knowledge to provide services for other similar products (competitors!) as well their own installed base.
    - Training and simulation: Complex products such as aircrafts, drones, moving machines and any tele-operated machineries need a simulation environment for proper training of the human driver/operator.
    - Condition monitoring, condition-based maintenance, anomaly detection, performance monitoring, prediction, management: The traditional service business sector is still encountering major challenges in practice. It will therefore require an extension to the above, as targets of services are expanded to other topics in customer businesses in addition to spare parts or condition monitoring.
    - Strategic partnership and data capitalization, "data as currency", will be key factors that will enable new business model inside AI service business
- **Fleet management, Edge and local/global decision making**
    - Decision and operations support: In most cases, decision-making is not automatic, whereas in the future it could be based on remote expert assistance or extensive diagnosing (AI-based, etc), engineering, and knowledge management systems.
    - Fleet management: This could benefit on the basis of sold items, by obtaining knowledge and experience from range of similar components and machines in similar or different conditions.
- **Business services integration**
    - Local and global services: Organising services locally close to customers and centrally at vendors' sites.
    - Full lifecycle tutoring: Monitoring activities, level of stress and performance-oriented behaviour during the product's life, from anticipating its end of life to properly handling its waste and recycling, including improved re-design for the next generation of products.

3.3.3.5 Major Challenge 5: Digital twins, mixed or augmented reality, telepresence

3.3.3.5.1 Status, vision and expected outcome

A "digital twin" is a dynamic digital representation of an industrial asset that enables companies to better understand and predict the performance and behaviour of their machines, and also to find new revenue streams. Currently, connectivity to the cloud allows an unprecedented potential for the large-scale implementation of digital twin technology for companies in various industries. A physical asset can have a virtual copy running in the digital twin environment, increasing revenue through continuous operational data. Similarly, several plants can operate in the same virtual space benefiting from globally shared data and remote automated decision-making. Factory/mill level digital twins are ever more autonomous and intelligent and can be adaptative by on-line learning. If needed they can interact to cloud level to evaluate model and optimize parameters using transfer learning at cloud.

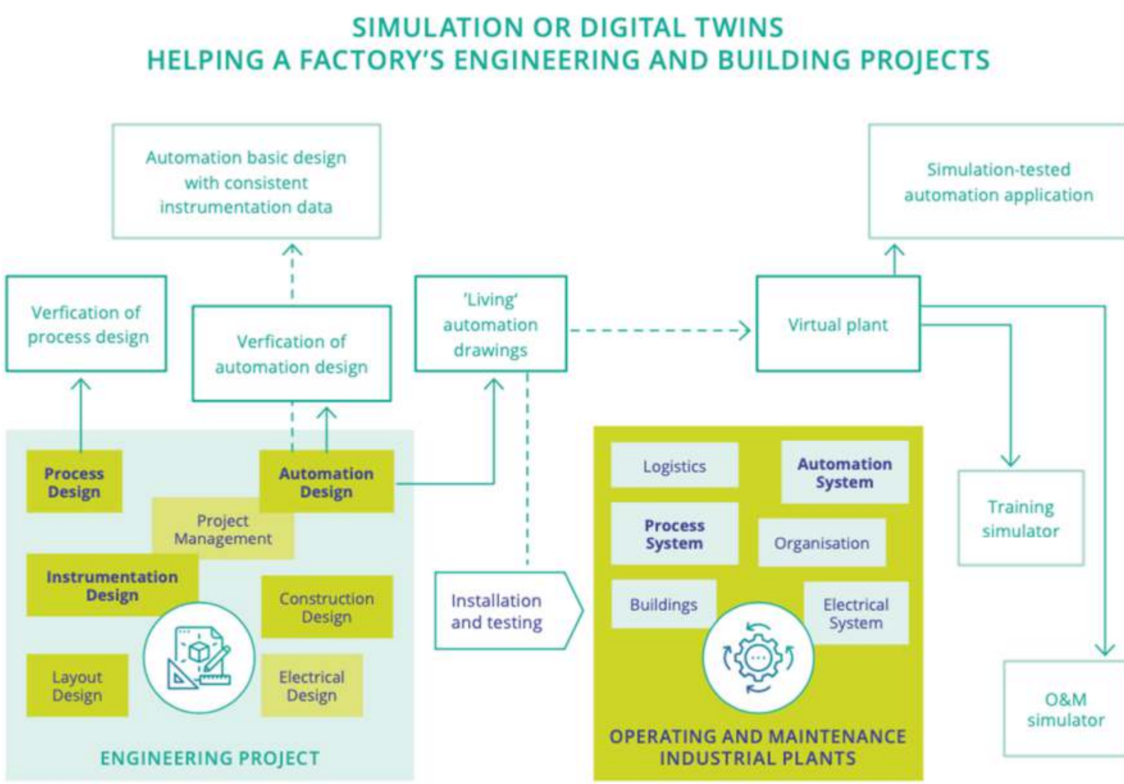


Figure 3.3.4 - Simulation or digital twins helping a factory's engineering and building projects

Simulation capability is currently a key element in the European machine tool industry's attempt to increase its competitiveness. In the Industry 4.0 paradigm, modelling plays a vital role in managing the increasing complexity of technological systems. A holistic engineering approach is therefore required to span the different technical disciplines and provide end-to-end engineering across the entire value chain.

In addition to virtual commissioning, modelling and simulation can more widely respond to many digitalisation challenges:

- visualising physical or real-world phenomena of products, production, businesses, markets, etc.
- helping designers to perform their core tasks – i.e., studying alternative designs, optimising solutions, ensuring safety, and providing testing for automation and Internet of Things (IoT) solutions.
- The effects of changes can be safely and more comprehensively assessed in advance in a virtual domain rather than using real plants, equipment or even mock-ups.
- Simulators offer versatile environments for users or operator training.
- It is evident that former computer-aided design (CAD)-driven digitalisation is shifting the focus towards simulation-based design.

Simulators may be used online and in parallel with its real counterpart to predict future behaviour and performance, provide early warnings, outline alternative scenarios for decision-making, etc., although they have years of research behind them, such tracking simulators are to be co-designed and improved exploiting also recent investments in computing infrastructures (e.g. HPC EU families, EPI initiative), with a special focus on the industrial context.

Telepresence technologies can also be considered as the predecessor for an extended reality (XR) presence. The combination of new and advanced technology like e.g. XR, Artificial Intelligence of Internet of Things (AIoT), Edge, High Performance Computing (HPC), 5G and open integration platforms offers significant potential for innovation, which would benefit the evolution of European digital industry.

As an example, XR is a combination of virtual and augmented reality, and an XR presence is a continuum between a physical reality presence and a virtual reality presence. The main driver here is improving competitiveness through better productivity, more effective worker safety and better quality. The industrial applications have followed the prospects offered by the gaming industry and consumer applications. One of the reasons for its increased take-up is the declining cost of electronic components and sensors.

As some major smart glass producers will provide technology and a platform for consumers, other EU industrial groups could do the same, and the EU industrial ecosystem could take benefit of such devices and provide industrial use cases. These can be extended then to state-of-art applications. “Back-end” services are Digital Twins and Condition Monitoring systems that will provide everything that optimizes the system work, providing critical and useful information for the Field Worker. Longer vision for the integration human actions and back-end data servers can be used to build knowledge graphs to help other users to work like experts. As a summary: smart glasses can extend human understanding and knowledge if services and information flow can be utilized and formed to usable knowledge.

The EC "Data Act" initiative[19] issued in 2024 promotes reliable and secure access to data within the European data economy and helps to advance the digital transformation by managing various aspects of data sharing, exchange and use. In terms of industrial applications and global interoperability between different legal entities, this means that data ownership and data use must be properly managed. For example, when digital twin processes data supplied by a tool manufacturer (data holder), then the user of this tool must be informed about this activity.

3.3.3.5.2 Key focus areas

- **Digital Twin: Design process digitalisation, telepresence**
  - Heterogeneity of systems: Information sharing and standards and means to ensure interoperability of digital twins and their information sources are important to facilitate information synchronisation. Having all relevant engineering disciplines (processes, assembly, electronics and electrical, information systems, etc.) evolving together and properly connected over the lifecycle phases is therefore crucial. This also involves multi-domain simulation, joint simulation of multi-simulation systems coupling.
  - Immersive telepresence for industrial robotics from design toward production lines and any other operational scope.
  - Digital twins applied to sustainability and circular economy: Simulate the usage of energy, use of raw material, waste production, etc., with the goal of improving energy efficiency and circular economy performance.
- **Virtual commissioning, interoperability:**
  - Virtual commissioning: Digital twins applied to virtual commissioning to bring collaboration between different disciplines and models from domains of engineering (mechanic, electronic, automation) in the same environment.
  - Interoperability: Applications cannot yet be used across platforms without interoperability.
- **Simulators: Tracking & Simulator based design**
  - Tracking mode simulation: Model adaption based on measurements. Generating simulators automatically from other design documentation, measurements, etc. Generation of simulators from 3D, data-driven models, etc.
  - Simulator-based design: Digital twin for testing the designed model by replacing the required physical components with their virtual models. This offers continuous design improvement (the digital twin provides feedback and knowledge gained from operational data), design optimisation, etc.
- **Digital twins combined with data-driven models (knowledge and data fusion):**.
  - Combination of data-driven and knowledge-based models along the complete lifecycle (product and production). The real challenge is to combine physics and

[19] https://digital-strategy.ec.europa.eu/en/policies/data-act

knowledge-based models (digital twins) with data- driven models (models created using AI from massive acquired experimental data), capitalising on the strength of information present in each of them.
    - New ways to generate large 3D scenes have been based, for example, on NeRF (Near Radial Field) and/or the Gaussian Splats.
- **Humans & Knowledge integration:**
    - Human-in-the-loop simulations: Methods and simulations for human-in-the-loop simulations and integration of digital twins in learning systems for workers.
    - The 3D Internet platform: to integrate all of the aforementioned aspects into a single powerful networked simulation for humans to get the Situational Awareness for an industrial process as a whole.
    - Live 3D Digital Twins: to provide the state awareness (animation & color etc.) in AR/VR/XR.

3.3.3.6 Major Challenge 6: Autonomous systems, robotics

3.3.3.6.1 Status, vision and expected outcome

Machines are usually more precise and efficient than humans when carrying out repeatable tasks. Thus, replacing or aiding work processes susceptible to human errors, quality defects and safety issues with machines will have an impact on quality and redundant waste. The application of AI in robotics and the same robotics without AI extends the opportunity for automation of manual tasks, increasing the sustainability, safety & security of the production and its green transition. This will reduce the environmental impact -efficient reduction of the waste and optimise product quality toward zero-defect, process, and manpower. This will also achieve a more safe and secure working area to ease the human-machine, and machine-machine co-working.

There are many kinds of autonomous systems, (co)-robots and working machines. Just to give an insight into their widespread adoption, can be categorised by purpose, as follows:

- Industrial machines and robots:
    - Manufacturing (e.g. welding, assembling, spray gun robots).
    - Material handling (e.g. conveyors, warehouse robots, trucks).
- Consumer robots:
    - Domestic (e.g. robotic lawn mowers or vacuum cleaners).
    - Care (e.g. lifting or carrying robots).
- Healthcare and medical robots:
    - Robotic surgery, hospital ward automation.
    - Medical tests and hospital care, remote healthcare.
    - Medical imaging, exoskeletons.
- Moving machines:

    - Mining machines (e.g. drilling machines, dumpers, conveyors).
    - Forestry (e.g. forest harvester), agriculture (e.g. tractors, appliances).
    - Construction (e.g. excavators, road graders, building robots).
    - Logistics and sorting centres (e.g. cranes, straddle carriers, reachers, conveyor belts, sorting machines, trucks).
    - Military robots and machines.
- Transport:
    - Vehicles, trucks and cars, trains, trams, buses, subways.
    - Aviation (e.g. aeroplanes, helicopters, unmanned aerial vehicles, UAVs).
    - Marine (e.g. vessels, ships, auto-piloted ships), submarine (e.g. auto-piloted submarines).
- Utilities and critical infrastructures:
    - Extraction (e.g. drills for gas, oil).
    - Surveillance (e.g. quadcopters, drones).
    - Safety, security (e.g. infrared sensors, fire alarms, border guards).
    - Energy power plants sensors and actuators (e.g. production and distribution).
    - Transportation (e.g. moving bridges, rail exchanges).

The main aims and evolution trends of robots and autonomous systems in digital industry are oriented toward:

- Production efficiency, speed and reduced costs,
- Higher precision and quality, safety in working conditions,
- To scale up “smart and high-end manufacturing”.

As is evident from the above, robots and machines and in particular autonomous systems are involved in all application chapters of this SRIA in addition to **Digital Industry** – i.e. Digital Society**,** Health and Wellbeing**,** Mobility**,** Energy**,** and Agrifood and Natural Resources- and are positively impacted by any improvement on both technological layers of the SRIA; foundational and cross-sectional technology.

There is undoubtedly a move to increase the level of automation and degree of digitalisation in industry, which will ultimately lead to fully autonomous systems.

However, between low and high technology manufacturing (two extremes: entirely manual and fully autonomous), there will always lie a large area of semi-autonomous equipment, units, machines, vehicles, lines, factories and sites that are worth keeping somewhat below 100% autonomous or digitised. The reasons for this include:

- A fully autonomous solution may simply be (technically) near to impossible to design, implement and test.
- If achievable, they may be too expensive to be realised.
- A fully autonomous solution may be too complex, brittle, unstable, unsafe, etc.
- A less-demanding semi-automatic solution may be easier to realise to a fully satisfactory level.

When the extent of automation and digitalisation are gradually, reasonably and professionally increased, often step by step, they may bring proportionally significant competitive advantages and savings that strengthen the position of digital industries overall. However, since the extent of automation and digitalisation remains well below 100%, any potential negative effects to employment are still either negligible or non-existent. On the contrary, the competitive advantages due to the adoption of robotics and autonomous system solutions increases the market positions of companies and, generally, enhances the need for more people in the respective businesses. An evolution of work organization and leadership may be required to enable the full potential of the digital industry.

3.3.3.6.2 Key focus areas

a) Autonomous functions of systems:

Advances in Artificial Intelligence simplify the development of fully autonomous systems, which can fully automate work tasks in various application domains by increasingly reducing the need for human intervention and, at the same time, allowing to highlight additional innovative functions. A prominent example can be found in the automotive domain, where autonomous vehicles are expected to play a key role in the future of urban transportation systems. Such a challenge in a so complex scenario will promote significant advances in any autonomous functions of the many systems which are integral part of the Digital Industry ecosystem.

Autonomous robotics is the key enabling technology for the implementation of autonomously functioning shopfloors. Immediate benefits will be additional safety, increased productivity, greater accessibility, better efficiency, and positive impact on the environment. The following figure provides a generalised view of technologies and functionalities used in autonomous systems developed in the automotive domain. Nevertheless, this schematic representation of required building blocks can be generalized to many other types of autonomous systems.

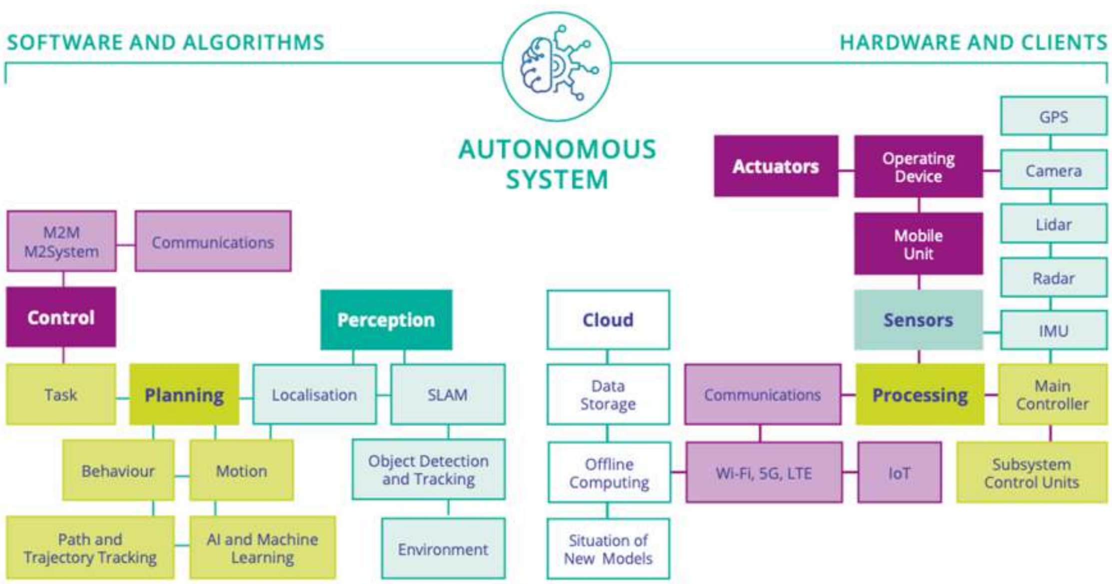


Figure 3.3.5 - A generalised overview of autonomous system (AS) technologies and functionalities. Adapted from Pendleton, S.D., Andersen, A., Du, X., Shen, X., Meghjani, M., Eng, Y.H., Rus, D., Ang, M.H.Jr. (2017). "Perception, Planning, Control, and Coordination for Autonomous Vehicles". Machines 5(1), p. 6.

The following figure shows matching of ISA95 standard that define control and other enterprise functions with building blocks of autonomous systems. The autonomy is expected to increase in the level 2 and Level 3 of ANSI/ISA95[20].

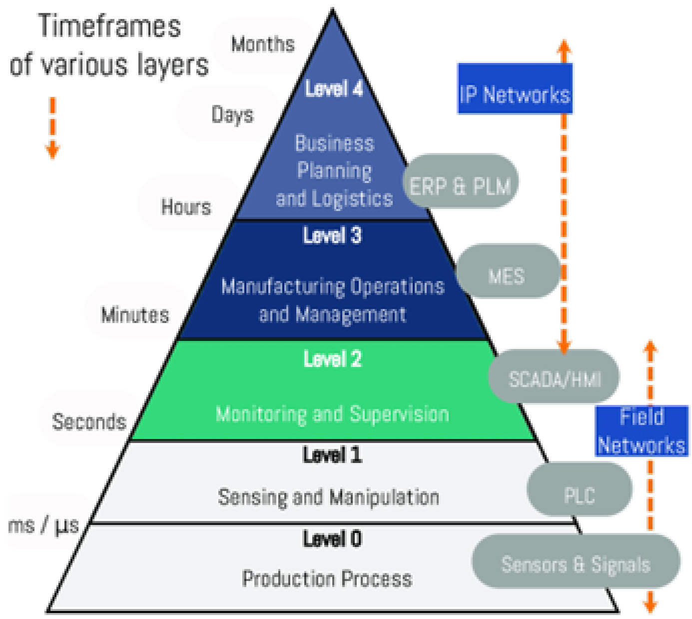


Figure 3.3.6 - ISA95 Hierarchy Model with building blocks of Autonomous Systems according to the 5 levels of the conventional automation pyramid[21]

b) Safety and security in autonomous systems:

Current standards of safety requirements for autonomous machines categorise safety into four approaches.

- On-board sensors and safety systems for machines that work among humans and other machines but is restricted to indoor applications.
- An isolated autonomous machine that works in a separated working area, mostly an intensive outdoor environment where other machines or humans are monitored.
- Machine perception and forecast of expected and/or unexpected human activities aimed at: (i) assisting human activities and movements with a proactive behaviour; (ii) preserving human health and safety; and (iii) preserving the integrity of machinery.
- An operator is responsible for reacting to a hazardous situation encountered by the autonomous machine when being provided with enough time between alert and transferring responsibility.

[20] https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa95

[21] https://www.researchgate.net/figure/The-conventional-automation-pyramid-according-to-the-ANSI-ISA-95-model-The-five_fig1_343473533

c) **Requirements management and conceptual modelling of autonomous systems:**

With the increasing complexity of autonomous functionality in both AV and ADAS systems, traditional methodologies of developing safety critical software are becoming inadequate. This observation can be made not only for autonomous driving, but in any industrial field of application of autonomous systems. Since autonomous systems are designed to operate in complicated real-world domains, they will be expected to handle and react appropriately a near endless variety of possible scenarios, meeting expectations from various stakeholders such as the internal engineering teams, people involved in the autonomous systems managements (e.g. passengers, drivers, workers, e-Health patients), regulatory authorities, and commercial autonomous vehicles/robots fleet operators.

d) **Human–machine interaction in autonomous systems:**

Improvements on sensing capabilities, on actuation control, IIoT (Industrial IoT) and SoS distributed capability are, in robotics and autonomous systems, the key enabler for:

- Human-robot interaction or human-machine cooperation.
- Transparency of operations between human and Advanced Machine Systems (AMS) in uncertain conditions.
- Remote operation and advanced perception, AS oversight and tactical awareness.
- Autonomy intended to enhance human capabilities.
- Natural human interaction with autonomous systems.
- Assisted, safety-oriented and proactive robot interaction with humans.

e) **Digital design practices including digital verification and validation (V&V):**

- Automatic or semi-automatic V&V.
- A digital design environment, digital twins, physical mock-ups.
- Sub-task automation development, generation of training data and testing solutions and field data augmentation, according to a handful of global machine manufacturers.
- Machine state estimation (assigning a value to an unknown system state variable based on measurements from that system).

f) **Simulators and autonomous systems:**

- Process model based and product 3D-models approaches, environment and object models, and simulation tools.
- Early design phase simulators.
- Robotic test environments.
- Empirical or semi-empirical simulators, making use of both real and simulated data collected from previous experiments.
- Off-road environments.

g) Autonomous capabilities development in a digital environment

- Autonomous decision taking.
- Self-evolving capabilities.
- Exploitation of knowledge in cognitive flexibility and in adaptability of the reaction.

3.3.4 REQUIREMENT OVERVIEW

The most obvious requirement for Digital Industry is the availability of all technologies, components and systems as described in other parts of this document both as foundation technologies and cross-sectorial areas serving and enabling the European Digital Industry mission.

It is extremely important that also a number of societal and policy needs are met, particularly the following aspects of EU legislation, related to working environments:

- Adoption of trustworthy, responsible AI, XR and robotics.
- To foresee exploitations of next generation HW architectures and new chip design (e.g. RISC-V, PIC).
- Adoption of any type of technology safeguarding safety and security of workers.

Also the following aspects related to EU policies are to be taken into consideration, for the promotion of:

- **Resilience** of EU production capabilities and supply chain towards Industrial EU sovereignty.
- **Sustainability** of EU manufacturing renovation and evolution towards a greener and safer EU.
- **Digitalisation** of EU Industry towards a quicker and better Innovation vocation, cost-efficient production, energy consumption saving and capacity to forecast market and societal needs.

3.3.5 TIMELINE

Major challenge	Topic	Short term (2025 – 2029)	Mid term (2030-2034)	Long term (2035 and beyond)
Major challenge 1: Responsive and smart production	**Topic 1.1:** Robust optimal production, scalable first time-right production	manufacturing plants more robust able to ensure continue production by means of the reacting capability to external/ internal disturbance	large adoption of self-healing and redundant automation systems, and capability of first time-right, zero-defect manufacturing	predictive maintenance empowering a very adaptable production
	Topic 1.2: · Mass customization and personalized manufacturing, customer-driven manufacturing	shop floor optimization on base of increased data availability due to larger adoption of digitalization and secure communication	lot-size-one manufacturing and personalized product design	greater production flexibility, Mass customization and personalized manufacturing, customer-driven manufacturing
	Topic 1.3: · Resilient and adaptive production, including the shortening of supply chains and modular and flexible factories	shortening of supply chains trend larger adoption of edge smart system/device in the shop floor	Adoption of even more modular and flexible lines/factories for a Resilient and adaptive production	Resilient and adaptive production
	Topic 1.4: · Cognitive production	larger integration of intelligent tools , cloud technology, embedded and remoted AI, complex analysis, big data, health monitoring of	Enabling new analytics and learning for responsive, sustainable and adaptive manufacturing involving natural and artificial cognition alongside IoT devices and data analytics	more cognitive production

	Topic 1.5: · Manufacturing as a service	production equipment and products, IoT infrastructures for ever more "as services" approaches within the manufacturing process overall chain.	Any technologically advanced solution capable of enlarging the geographical distribution of production and strengthening the split and outsourcing of production phases within the manufacturing supply chain	More Manufacturing as a service, moveable factories, more resilient supply chain ecosystem
	Topic 1.6: · · Embedded/Edge/Cloud architectures		larger adoption of 5G private networks nodes equipped with very powerful computing technology; Edge-to-Cloud and edge-to-edge communication and updating capability, embedded AI	Workload and services moved from the centralized data centres (core network) ever more to the proximity of production lines
	Topic 1.7 · ·Standardization	continuous improvement of the integration and compatibility of the different standards in use		
Major challenge 2: Sustainable production	**Topic 2.1:** · Monitoring flows of energy, materials, waste and Lifecycle assessment	larger adoption of edge smart system/device and embedded AI in the production line	Sustainable production driven by those industrial sectors that need to fully track material and energy flows (food, medicine, etc.)	
	Topic 2.2: · Virtual AI assistants	Gen AI assistant and AI optimizer to help operators locally by providing advice based on LLM	Reducing the impact of human factor errors in the entire production chain (from design to production bay)	
	Topic 2.3: · Human–machine interfaces	Wearable and foldable devices; Gesture, eye and	Augmented reality; decision-making support by	Improved human-machine and machine to machine interfacing

	and machine-to-machine communications	natural language command recognition, A&VR to become part of Advanced Human Machine Interface	enhanced visualization of data and analytic results	
	Topic 2.4: · Human operators in more autonomous plants and in remote operations	development of robotic technologies (haptic, dexterity, active and return of force exoskeletons…) and system solutions for advancement on human machine cooperation and teleoperation in intuitive and safe manner	integration of Localization devices for indoor/outdoor navigation and position tracking	coordinated and distributed system solutions, able to ensure a smart and safe workspace where humans collaborate with machines using the strengths of both simultaneously
	Topic 2.5: A· Human safety	localization of personnel, machines and vehicles, situation-aware safety	sensing of safety issues, proximity detection, online human risk evaluation, map generation, enabling an assistive control for co-working interaction	
	Topic 2.6: · Competence and quality of work in a human-centered manufacturing	ability to attract skilled personnel and to maintain their competence over time		
	Topic 2.7: · Green deal	high efficiency, low energy usage, carbon-neutrality, or zero-carbon usage, zero waste from water, soil and air-measured, calculated or estimated on product, factory, global and lifecycle levels.		zero global warming emissions by 2050 Recycling, re-use, and de- and re-manufacturing
Major challenge 3: Artificial Intelligence in digital industry	**Topic 3.1:** European AI framework	Set out the context for the operation of AI public/private partnerships (PPPs), as well as other PPPs or joint undertakings (JUs). Define the primary areas of importance for AI research, innovation, and deployment. Consider and evaluate the impact of AI on stakeholders and users of the products and services by means of initiatives on AI.		
	Topic 3.2: AI in manufacturing	AI for dynamic production planning and management Virtual models spanning all levels of the factory life and its		Improving process integration at the shop-floor level to

		lifecycle AI for green/sustainable manufacturing AI applied in supply chain management AI for advanced manufacturing processes AI for adaptive and smart manufacturing devices, components, and machines.	increase flexibility and dynamic production path on base of machine workloads, inventory, and production plans with a tightly coordination (vertical integration) with higher-level business processes such as procurement, quality control, marketing, sales, and distribution
	Topic 3.3: AI for decision-makings	explainable AI to increase "human-centric" approaches in autonomous industry	AI to support complex decision-making (Human, machine, mixed decision-making) processes
	Topic 3.4:AI for monitoring and control	AI for control technologies. AI for monitoring services. AI for maintenance systems for increased reliability of production systems. AI services for continuous evaluation and mitigation of manufacturing risks. AI for quality inspection.	to exploit industry environment digitalization in an adaptive control to achieve smart horizontal integration (i.e. dynamic path of manufacturing through production stations on base of machine workload etc.)
Major challenge 4: Industrial service business, lifecycles, remote operations, and teleoperation.	**Topic 4.1:** · Remote operations, teleoperation	conventional industry infrastructures have to be adequate to protect from cyber threats and validated for new advanced applications ensuring continuous risk assessment and monitoring	Operating or assisting in operations of industrial systems from remote sites. Implement distributed service applications on effective edge cloud systems. Large adoption of telecommunication network (5G) with very low latency for remote operations
	Topic 4.2: · AI Services for monitoring	Collaborative product-service engineering, lifecycle engineering	Exploit AI ability to recognize a strategy, an outline from highly

	and collaboration	Training and simulation Condition monitoring, condition-based maintenance, anomaly detection, performance monitoring, prediction, management new business model inside AI service business based on data capitalization		variable patterns, in the definition of the state and maintenance of the machineries, in the optimization of the abstract models, in the identification and in the definition of new control strategies by means secure exchange of data
	Topic 4.3: · Fleet management, Edge and local/global decision making	ever more unmanned vehicles\platforms, integrated in the factory layout to provide manufacturing needs and support production paths	Fleet management: Decision and operations support also based on remote expert assistance capitalize in services knowledge and experience from range of similar components and machines in similar or different conditions	enable an assistive control for co-working interaction improve customer remote assistance
	Topic 4.4: ·Business services integration	Exploit the interoperability between different production frameworks and other engineering tools in the field of production planning and modelling	Ai support on organizing services locally close to customers and centrally at vendors' sites. AI support on full lifecycle tutoring	to improve customer services and technical support, and to make customer active in product definition and delivery
Major challenge 5: Digital twins, mixed or augmented reality, telepresence	**Topic 5.1:** · Digital Twin: Design process digitalisation, telepresence	Thing Description as abstraction of physical entities for advanced interactions exploiting virtualization potentialities	Immersive telepresence for industrial robotics from design toward production lines and any other operational scope. Digital twins applied to sustainability and circular economy Ensure interoperability of digital twins among systems' Heterogeneity	
	Topic 5.2: Virtual commissioning,	strength collaboration of different functional teams on the	virtual commissioning to bring collaboration between different engineering domains on same environment. Platforms	

	interoperability	same common objective/business model	interoperability	
	Topic 5.3: Simulators: Tracking & Simulator based design	Model adaption based on measurements. Generating simulators automatically from other design documentation, measurement Digital twin for testing the designed model by replacing the required physical components with their virtual models		to support the design and development of completely new plants or new production in an existing plant
	Topic 5.4: Digital twins combined with data-driven models (knowledge and data fusion	combine physics and knowledge-based models (digital twins) with data- driven models (models created using AI), capitalizing on the strength of information present in each of them New ways to generate large 3D scenes		
	Topic 5.5:· Humans & Knowledge integration	Human-in-the-loop simulations: The 3D Internet platform Live 3D Digital Twins: to provide the state awareness in AR/VR/XR	Allow a collaborative workspace, coordinated and distributed system solutions able to ensure a smart and safe workspace where humans collaborate with machines using the strengths of both simultaneously	
Major challenge 6: Autonomous systems, collaborative robotics	**Topic 6.1:** Autonomous functions of systems	Advancements in Artificial intelligence to easier achieve fully autonomous systems solutions human intervention reduction, introduction of additional innovative functions Autonomous robots enabling implementation of autonomous functions shopfloors Advancement on energy autonomous node (energy harvesting in sensors actuation)		Self-learning and control systems in increasing advanced functionalities and smartness and awareness within the workspace Automated and innovative data-based services to reduce the offline measurement, increase fab automation reducing production cost and times
	Topic 6.2: Safety and security in	Automated vision monitoring & Machine and robot vision in	assisting human activities and movements with a	Collaborative workspace, coordinated and distributed system

	autonomous systems	support to machine perception capabilities and to forecast of expected and/or unexpected human activities	proactive behavior; Preserving human health and safety; Preserving the integrity of machinery	solutions able to ensure a smart and safe workspace where humans collaborate with machines using the strengths of both simultaneously
	Topic 6.3: Requirements management and conceptual modelling of autonomous systems	involvement in autonomous systems definition of the expectations from various stakeholders such as the internal engineering teams, people involved in the autonomous systems managements (e.g. passengers, drivers, workers, e-Health patients), regulatory authorities, and commercial autonomous vehicles/robots fleet operators.		Increased complexity of autonomous functionality autonomous systems able to handle and react appropriately a near endless variety of possible scenarios
	Topic 6.4: Human–machine interaction in autonomous systems	worker's safety coordination by security of data exchange and low latency communication Improvements on sensing capabilities, on actuation control, IIoT and SoS distributed capability	Human-robot interaction (HRI): supporting the development of robotic technologies (haptic, dexterity, active and return of force exoskeletons…) and system solutions for advancement on human machine cooperation and teleoperation in intuitive and safe manner. Also support the intuitive programming of robots (e.g. voice/gesture control)	
	Topic 6.5: Digital design practices including digital verification and validation (V&V)	Process improvements based on shared information allowing interaction in the value chain from product design up to deliver; Automatic or semi-automatic V&V. A digital design environment, digital twins, physical mock-ups.	Sub-task automation development, generation of training data and testing solutions and field data augmentation, according to a handful of global machine manufacturers. Machine state estimation	Supporting design of product and process development in manufacturing phases providing a fast settling of production lines and to improve batch production and Lot size"1 capability" and mass customization
	Topic 6.6: Simulators and autonomous	to support the design and development of completely new	Process model based and product 3D-models	Empirical or semi-empirical simulators, making use of both

	systems	plants or new production in an existing plant by simulation on virtual plant to better prepare physical production and verify product-production constraints	approaches, environment and object models, and simulation tools. Early design phase simulators. Robotic test environments.	real and simulated data collected from previous experiments. Off-road environments
	Topic 6.7: Autonomous capabilities development in a digital environment	Distributed processing capability on base of the robustness of communication and low latency on answer	Autonomous decision taking. Self-evolving capabilities. Exploitation of knowledge in cognitive flexibility and in adaptability of the reaction.	

3.4

ECS Key Application Areas

HEALTH AND WELLBEING

3.4. HEALTH AND WELLBEING

3.4.1. SCOPE

Technological effervescence

In the field of Digital Health, a significant discrepancy can be observed. There are thousands of digital health solutions available on the market, and thousands more are being developed, mostly by tech companies (including start-ups) that are entering the sector for the first time, rather than traditional healthcare companies. Many of these companies have high expectations, but they are learning that these expectations can be challenging to achieve in such a complex and highly regulated sector. As a result, only a small fraction of Digital Health innovations succeeds in the market. Certain aspects of the healthcare industry may even hinder market entry, especially concerns regarding the reimbursement of medical technologies.

Figure 3.4.1 - Harvard Business Review: "Why innovation in healthcare is so hard?"

Another issue lies in the conventional practices regarding the manufacturing of medical devices, considering that medical devices are developed in this context as a dedicated instrument for a specific application, limiting the expected return-on-investment. However, gradually, change is coming. The objective of the Chapter is to highlight the business opportunity that the current acceleration of Health Digital Transformation should bring for the ECS community.

Healthcare systems sustainability

Current megatrends highlight the ubiquity of healthcare in our society. Urbanisation is directly correlated to population density with healthcare associated issues. The ageing demographic means more health issues like chronic diseases. The COVID 19 pandemics has shown that specific health situations may disrupt access to health services at global level, while it already needed to be improved for several categories of the world population.

Healthcare costs are escalating worldwide. In 2040, the world will be collectively spending an estimated $20 trillion every year on healthcare, representing a 150% increase compared with 2014, according to a study published in The Lancet. This explains why almost all countries have established an explicit national Digital Health strategy associated with a roadmap to deploy solutions. Technology adoption, through the healthcare digital transformation, is considered as a pivotal element to secure the sustainability of health systems worldwide.

Opportunity for the European Electronic & Components System industry

The emergence of so-called P4 healthcare – predictive, preventive, personalised, participatory – as opposed to reactive healthcare, is blurring the lines between medical technology (MedTech), the pharmaceutical industry and electronic components and systems (ECS) industry, opening the way for healthcare innovation. This is a huge opportunity for the European ECS industry, its worldwide MedTech and pharma market leaders, as well as the 25,000 SME MedTech companies across Europe, since the new healthcare ecosystem will rely on digital instruments, advanced electronic sensors and photonics, microelectromechanical systems (MEMS), and the large volume, high-quality, low-cost production capabilities of the ECS industry.

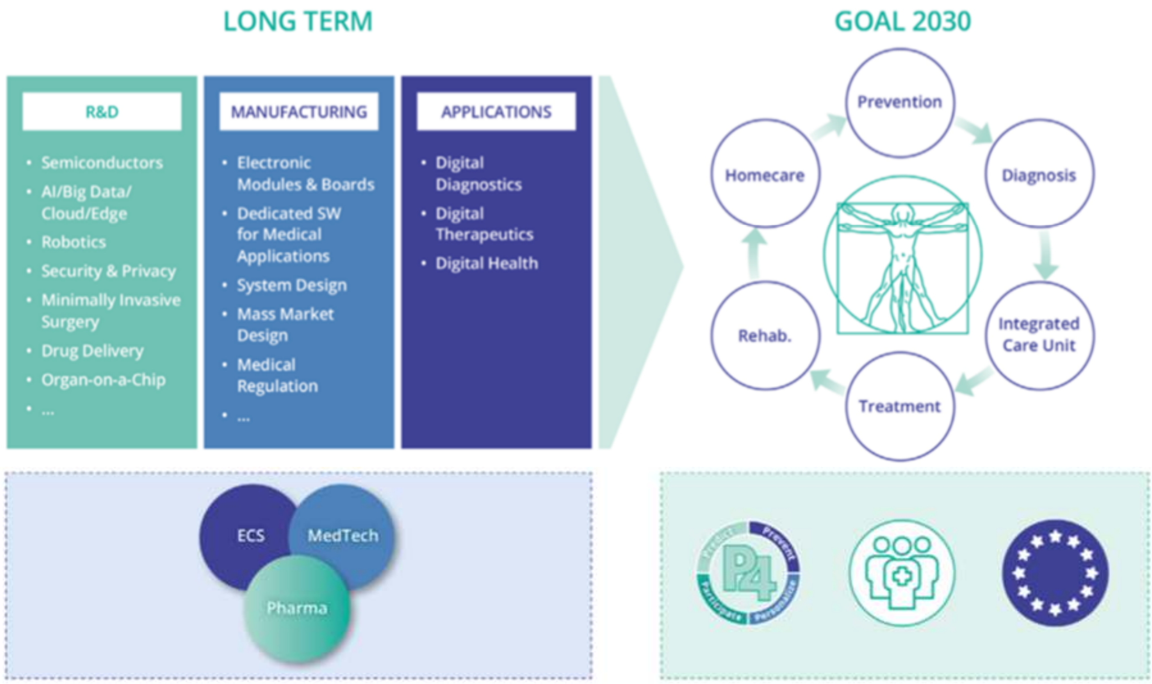


Figure 3.4.2- ECS Industry impact in Healthcare Digital Transformation

3.4.2. APPLICATION TRENDS AND SOCIETAL BENEFITS

3.4.2.1. Application trends

Digital health is a relatively new field that is still in its early stages of development. It is based on aspects like wearable devices, remote patient monitoring, prevention, ageing well and P4 medicine which altogether are showing great promise for health and wellbeing. The technologies enabled by the chip industry have the potential to significantly disrupt the healthcare industry. They will change the way healthcare is delivered, make healthcare more accessible, and improve patient outcomes. Technical solutions shall also cover the full health value chain, from prevention, diagnosis, acute treatment to rehabilitation.

First, innovation in healthcare is connected to “Industry 4.0”. “Healthcare 4.0" and "Industry 4.0" are related concepts influenced by different factors specific to their respective industries. Healthcare 4.0 refers to the application of Industry 4.0 technologies specifically within the healthcare sector. The adoption rate of these technologies is slower in health and care due to different regulations, ethical considerations, and infrastructure.

“Healthcare 4.0” technologies are fundamentally changing the way healthcare is delivered and received, leading to a more connected, personalised, and data-driven health landscape leading to “integrated care” e.g., the integration of various components of healthcare facilitating coordinated care, multidisciplinary collaboration, personalised medicine...

“Healthcare 4.0” means moving away from centralised clinical models, with a greater focus on patient care outside of hospitals and on behavioural health. “Healthcare 4.0” also has the potential to break down the fragmentation between disciplines. Remote patient monitoring is for instance a cross-cutting application to different pathologies and use cases and opens broader perspectives for manufacturers than the dedicated medical device for a specific pathology. Another correlated and important aspect amplifies this “platform effect”. The boundaries between Pharma, MedTech and electronics companies are blurring due to the digitisation of healthcare, e.g., “Healthcare 4.0”. MedTech’s are focusing on the development of advanced wearable devices, implantable sensors, diagnostic tools, and connected healthcare solutions to improve patient monitoring, diagnostics, treatment, and overall healthcare delivery. Pharmaceutical companies are leveraging digital health solutions, such as mobile applications, wearable devices, and remote monitoring systems, to enhance patient engagement, monitor treatment adherence, collect real-time health data, and provide personalised support.

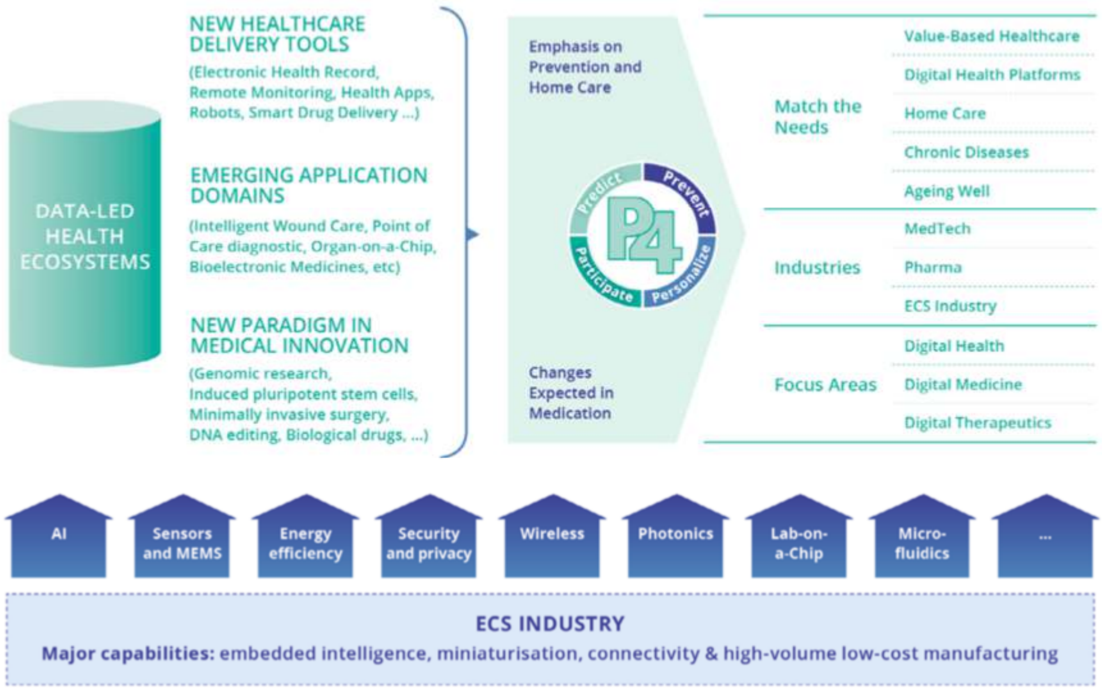


Figure 3.4.3 - Synergies in Data-Powered Health: Unveiling P4 Medicine & Future Use Cases through Industry Convergence (Source: STMicroelectronics)

Second, the consumerisation of medical devices leads to an increased demand for and usage of microelectronic components and systems. The consumerisation of day-to-day health is tech-driven with both the involvement of "Big Tech" (Apple, Google, Amazon...) and its most obvious illustration under the form of wearables (watches, hearables, rings, glasses, ...) and consumer apps. This is changing the way "Health Consumers" or even "Patients" expect to interact with healthcare.

Figure 3.4.4 -Healthcare consumerisation: GAFAM's Preponderant Role in Reshaping the Future of Health Tech

Consumerisation is essential for healthcare systems sustainability, as it emphasises preventive care and early intervention. Consumerisation also encourages innovation and disruptive technologies within healthcare, enabling new care models where patients actively participate in treatment decisions and care plans. Digitisation of healthcare systems is severely lagging the consumer domain and needs reinforcement due to its 'under-the-hood' nature and high complexity and custom requirements, further complicated due to the typically, small product volumes. Consequently, all benefits that we know so well from the consumer domain are missing with many healthcare systems.

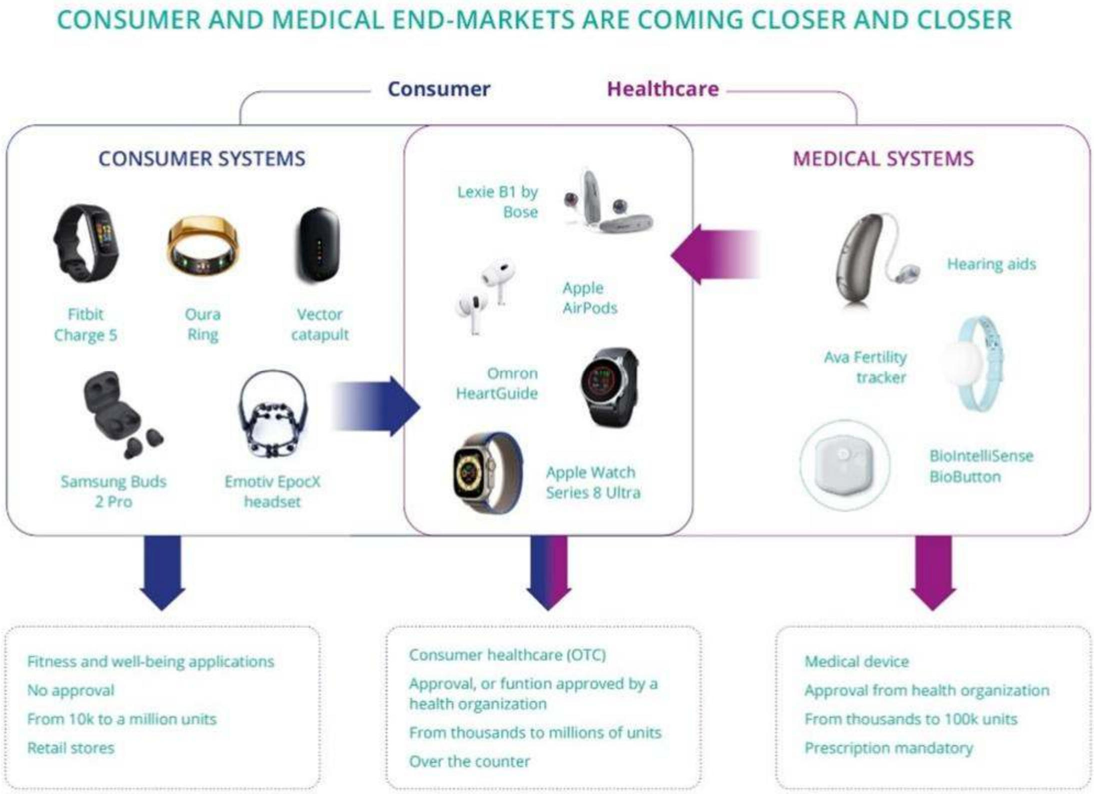


Figure 3.4.5 - The Merging Paths of Consumer and Medical domains (source: Yole Group)

Third, data makes customisation on a large scale possible. Customisation is the driving force behind the trend towards P4 medicine. Mass customisation combines the advantages of mass production (high volume, efficiency, and cost-effectiveness) with the benefits of customisation. Mass customisation in healthcare is closely linked to wearable devices (watches, hearables, rings, glasses, …) and the Internet of Things (IoT), requires AI and machine learning, allows personalised treatment plans through smart drug delivery systems, and may involve point-of-care diagnostic solutions that enable rapid and personalised testing.

3.4.2.2. Societal benefits

Companies in the MedTech and Pharma industries are transitioning from focusing on individual therapies to providing all-inclusive healthcare solutions, develop scalable platforms that can host a range of digital health services utilising smart devices and data analysis to enhance therapy outcomes and align with value-based healthcare principles. These initiatives will improve the general well-being and quality of life for many people. As far as diseases like dementia, hearing and vision impairment is concerned stigmatisation and isolation can be mitigated and allow patients to expand their active participation in social life.

Beyond increased accessibility and cost efficiency, incorporating artificial intelligence and advanced data analytics facilitates better collaboration among healthcare professionals to make highly informed decisions, resulting in enhanced patient outcomes. Moreover, the implementation of predictive analytics aids in the timely identification of diseases, promoting early intervention and preventing further complications. Patients have a more interactive and engaged healthcare experience.

Hospitals and clinics can use data analytics to predict patient inflow and manage their resources more efficiently. Big data analytics can provide valuable insights into disease patterns, treatment outcomes, and patient behaviours, driving research and innovation, and also monitor and respond to global health crises, like pandemics.

A well-known issue in the Digital Health transformation is that regulation must keep pace with innovation. Innovation is potentially limited by a lack of regulations, specifically in innovative areas such as mobile health, digital therapies, and technologies such as big data or AI. In the European landscape, regulatory hurdles, ecosystem transformation, accurate and redesigned payment schemes are currently addressed to allow the EU to take a strategic advantage in the Healthcare digital transformation.

3.4.3. MAJOR CHALLENGES

Five Major Challenges have been identified for the healthcare and wellbeing domain:

- **Major Challenge 1**: Enable digital health platforms based upon P4 healthcare.
- **Major Challenge 2**: Enable the shift to value-based healthcare, enhancing access to 4P's game-changing technologies.
- **Major Challenge 3**: Support the development of the home as the central location of the patient, building a more integrated care delivery system.
- **Major Challenge 4**: Enhance access to personalised and participative treatments for chronic and lifestyle-related diseases.
- **Major Challenge 5**: Ensure more healthy life years for an ageing population.

3.4.3.1. Major Challenge 1: Enable digital health platforms based upon P4 healthcare

3.4.3.1.1. Status

The MedTech industry is in the process of transitioning from an industry primarily producing high-end hospital equipment to one that will increasingly serve point-of-care (PoC) professionals and "health consumers", thereby moving from a product-based approach to the provision of "integrated services".

Electronic medical technology, such as the Internet of Medical Things (IoMT), minimally invasive implants, energy-efficient devices, advanced analytics, cognitive computing for advanced clinical decision support, cybersecurity, enhanced network capabilities for continuous data access (to mention only a few of those listed below in the section "Key focus areas"), will support the deployment of P4 healthcare in a data-led environment.

3.4.3.1.2. Vision and expected outcome

The P4 healthcare vision is not only placing doctors and other health professionals at the centre of the care process, but all those relevant to the health consumer. Even if the healthcare ecosystem is operating in a highly regulated environment, by 2030 we can expect this trend to progressively become the norm. The ECS community should participate in the development of dynamic healthcare systems that learn in real time from every result, achieve a better understanding of treatment response and prognostic heterogeneity, and introduce more refined, patient-tailored approaches to disease detection, prevention, and treatment.

The P4 healthcare vision, enabling for instance early diagnostics based upon merged data and machine-learning techniques through the detection of weak signals, allows preventive treatments that are far less intensive than acute treatments and increase the chances of survival and quality of life. The MedTech industry is not alone on this journey. New pharmaceuticals and treatments will be developed for personalised medicine settings by embedding connected devices and exploiting the potential of the IoT and AI.

This is all creating a new industry, one that revolves around digital health platforms. This platform-based new-market disruption will enable the emergence of specialised platforms, and new players will enter the health domain. This will impact current business models in healthcare, using aggregated data to create value rather than devices – supporting, for instance, proactive services, facilitating outcome evaluation for the treatment of different therapies, and paving the way for outcome-based or pay-per-use reimbursements. This is a potential path to reducing the burden of healthcare expenditures.

3.4.3.1.3. Key focus areas

The addition of AI capabilities – person-centred AI-based consumer devices/embedded AI-based medical devices and systems – to Electronic Components and Systems will significantly enhance their functionality and usefulness, especially when the full power of such networked devices is harnessed – a trend that is often called "edge AI". AI enables much more efficient end-to-end solutions by switching from a centralised to a distributed intelligence system, where some of the analysis carried out in the cloud is moved closer to the sensing and actions.

This distributed approach significantly reduces both the required bandwidth for data transfer and the processing capabilities of cloud servers. It also offers data privacy advantages, as personal source data is pre-analysed and provided to service providers with a higher level of interpretation. It also offers greater reliability and safety.

A high level of digital trust – for privacy and security by design, hardened and embedded AI models – is of course required for executing transactions in healthcare and wellbeing. Securing the IoT ecosystem is a multiple level problem. Privacy should be "by design". In general, integrating security features into an existing system can become very complex, sometimes impossible, and often increases the cost of the final product significantly. A more efficient approach is to consider those security requirements at the very beginning of a project, and then integrate them in the design and development phase. The ECS industry can assist with end-to-end solutions by providing on-chip security, supplying comprehensive hardware and software services, including authentication, data encryption and access management.

Next-generation connectivity – better performing, more ubiquitous, accessible, secure, and energy-efficient networks – will contribute to unleashing the potential of digital health. One of the main characteristics of future networks will be their increased intelligence to improve the performance of the networks, and offer sophisticated and advanced services to the users, due to edge computing and metadata, for instance.

With the significantly growing number of wearables and other battery-operated devices with small form-factor, very low power consumption is a major technology challenge for product designers. The transition from linear to circular economy will require innovative designs for the lifetime of electronic components and systems, and disruptive changes in ECS supply chains, to reduce the ecological footprint. The ECS industry will contribute to improving energy efficiency – including new, sustainable, and biocompatible energy harvesting – to locally process data and the transmission of pre-processed data as opposed to the transmission of high-volume data (such as imaging data).

As a result of improved integration and analysis of multimodal data, new tools for clinical decision-making and precision medicine will emerge, supporting early diagnostics, personalised medicine and potential curative technologies (e.g. regenerative medicine, immunotherapy for cancer).

3.4.3.2. Major Challenge 2: Enable the shift to value-based healthcare, enhancing access to 4P's game-changing technologies

3.4.3.2.1. Status

A major trend in healthcare is the transformation of large healthcare systems to an optimised hospital workflow: a shift from general hospitals treating any diseases towards integrated practice units that specialise in specified disease types. These units, organised around a medical condition, aim to maximise the patient's overall outcomes as efficiently as possible, increasingly through remote access, for patients anywhere in the world.

3.4.3.2.2. Vision and expected outcome

Pay-for-cure rather than pay-for-treatment can be an effective way to increase the efficiency of healthcare by avoiding unnecessary tests, therapies, and prescriptions. Combined with empowered patients, care-givers should be able to make better informed and more effective choices for treatment. To achieve this, outcomes need to cover the full cycle of care for the condition and track the patient's health status aftercare is completed. This first involves the health status, relying for instance on EHRs supporting precise communication between different care-givers' PoC diagnostic systems or AI-based clinical decisions. Early diagnosis is key for the successful treatment of both modest and challenging medical conditions.

Health outcomes are also related to the recovery and the sustainability of health. Readmission rates, level of discomfort during care, and return to normal activities should be taken into account for both providers and patients. Humanoid robots applying interpreted human body language and emotion in care delivery, sensors, the deployment of companion devices anticipating and contextualising acute or chronic conditions in EHRs involving health models describing the outcome health values for the patients, both in the short term and long term, will have a direct positive effect on readmission prevention.

To achieve this transformation, a supporting IT platform is necessary. Historically, healthcare IT systems have been siloed by department, location, type of service, and type of data (for instance, diagnostic imaging). An innovative and efficient healthcare information infrastructure – integrating IoT with big data learning for optimising workflow, usage, capabilities and maintenance, and of course digital trust – will aggregate the different areas for efficient value-based healthcare, combining prevention tools, early detection, and treatment. This will enable better measurement and facilitate the design and implementation of new bundle-based reimbursement schemes, reducing costs while improving health outcomes. By 2030, value-based healthcare will enable the adoption of optimisation practices already supported by ECS technologies in the industry.

3.4.3.2.3. Key focus areas

By 2030, clinical decision-making will be augmented by electronic medical records. Digital centres will enable advanced capabilities for clinical decision-making where AI, real-time data from portable Point-of-Care devices, 3D printing for surgeries, continuous clinical monitoring – including robotics to improve treatments either in the operating room, minimal invasively inside the body, at the general practitioner or at home – will support the integration of specialised care units. Many images will be combined with other sensor data and biomedical models to obtain precise, quantified information about the person's health condition, preventing and providing, for instance, early warnings for (combined) diseases supported by patient health models on complex health conditions. Low–latency, massive image processing is a major information source for AI-based automation, visualisation, and decision support within the whole care cycle. Precise quantified and annotated imaging is needed at many levels: from molecular imaging up to whole body imaging. The development and use of accurate digital twins of the human body will enable clinical trials, individualised computer simulations used in the development or regulatory evaluation of a medicinal product, device, or intervention. While completely simulated clinical trials are not yet feasible, their development is expected to have major benefits over current in vivo clinical trials, which will drive further research on the subject. Moreover, "digital twins" will help in combining all the data on a personal level and enable personalised clinical decision-making. Full electronic medical records will also enable optimized hospital workflows that increase efficiency and therefore reduce burden on medical practitioners.

Europe is a leading producer of diagnostic imaging equipment. In diagnostic imaging, the ECS industry has begun to place great emphasis on accurate radiation dose monitoring and tracking. Healthcare providers are already applying dose management as part of the quality programme in their radiology departments, and patient-specific computed tomography (CT) imaging and personalisation of scan protocols will be a key aspect of patient-centred care in radiology departments, facilitating the management and control of both image quality and dose with the optimisation of 3D X-ray imaging protocols. In particular, CT and X-ray systems based on the emerging spectral X-ray photon counting technology will enable new ways of imaging in order to reduce dose, improve tissue differentiation, enhance visualisation, or quantify materials. Image analysis and semi-automated decision making including applications like digital pathology will significantly benefit from underlying AI based methods.

The enormous capabilities of the ECS industry in miniaturisation, integration, embedded intelligence, communication and sensing will have a major impact on the next generation of smart minimally invasive devices:

- Highly miniaturised electrical and optical systems realised using advanced cost-effective platform technologies will bring extensive imaging and sensing capabilities to these devices, and enable the second minimally invasive surgery revolution, with smart minimally invasive catheters and laparoscopic instruments for faster and more effective interventions.
- Sensing and diagnostics solutions need to achieve appropriate sensitivity, specificity, and time-to-result.
- Reducing waste is possible through sensors made of biological materials, combining a biological component with a physicochemical detector.
- Fusion of diagnostics and surveillance will help reducing system and operational costs.
- To realise next-generation smart catheters, a broad spectrum of advanced ECS capabilities will need to be brought together, foremost in dedicated platforms for heterogeneous miniaturisation and integrated photonics. These can be complemented with platforms for embedded ultrasound, low-power edge computing, and AI and digital health platform integration.
- Optical coherence tomography (OCT) is another example where ECS technologies make a critical impact, in shrinking devices and reducing costs, allowing devices to be used in wider fields beyond ophthalmology.

Finally, it should be noted that the development of the next generation of smart minimally invasive instruments will go hand in hand with the development of new navigation techniques:

- Breakthrough innovations in photonics are enabling optical shape-sensing techniques that can reconstruct the shape of a catheter over its entire length.
- MEMS ultrasound technology will enable segmented large-area body conformal ultrasound transducers that are capable of imaging large parts of the body without the need for a trained sonographer, to guide surgeons in a multitude of minimally invasive interventions.
- Combined with other technologies, such as flexible and conformal electronics, low power edge computing, AI, and data integration into clinical systems, new optical and

acoustic- based technologies may eliminate the use of x-rays during both diagnosis and interventions, enabling in-body guidance without radiation.
- Augmented reality can be used for image-guided minimally invasive therapies providing intuitive visualisation.

As mentioned above, outcomes should cover the full cycle of care for the condition and track the patient's health status once care has been completed. Biomarkers derived from medical images will inform on disease detection, characterisation, and treatment response. Quantitative imaging biomarkers will have the potential to provide objective decision-support tools in the management pathway of patients. The ECS industry has the potential to improve the understanding of measurement variability, while systems for data acquisition and analysis need to be harmonised before quantitative imaging measurements can be used to drive clinical decisions.

Early diagnosis through PoC diagnostic systems represents a continuously expanding emerging domain based on two simple concepts: perform frequent but accurate medical tests; and perform them closer to the patient's home. Both approaches lead to improved diagnostic efficiency and a considerable reduction in diagnostic costs:
- Point-of-care testing (PoCT) methodology encompasses different approaches, from the self-monitoring of glucose or pregnancy, to testing infectious diseases or cardiac problems. However, it should be remembered that disposable PoC devices will need to be environmentally friendly in terms of plastic degradation and the replacement of potentially harmful chemicals.
- The key enabling components of current PoCT devices must include smart and friendly interfaces, biosensors, controllers, and communication systems, as well as data processing and storage.
- The emerging lab-on-a-chip (LoC) solutions, embedding multiple sensor platforms, microfluidics, and simple processing/storage elements, are currently the most promising basis for the realisation and development of accurate, versatile, and friendly portable and wearable PoCT devices. Their simplified operation mode eliminates the constraint of molecular biology expertise to perform a real-time reverse transcription polymerase chain reaction (RT-PCR) test, will enable innovative in vitro diagnostic (IVD) platforms, making possible decentralisation from highly specialised clinical laboratories to any hospital lab and near-patient sites, with dedicated sample prep cartridges, a more efficient prevention (referring to the recent Covid-19 pandemic) and prompt personalised diagnosis.

In addition, digital supply chains, automation, robotics, and next-generation interoperability can drive operations management and back-office efficiencies. Using robotics to automate hospital ancillary and back-office services can generate considerable cost and time efficiencies, and improve reliability. Robotic process automation (RPA) and AI can allow care-givers to spend more time providing care. For instance, robots can deliver medications, transport blood samples, collect diagnostic results, and schedule linen and food deliveries – either as a prescheduled task or a real-time request. Robotic processes also can be used for certain hospital revenue cycle and accounting/finance functions, such as scheduling and claims processing.

3.4.3.3. Major Challenge 3: Support the development of the home as the central location of the patient, building a more integrated care delivery system

3.4.3.3.1. Status

The trend towards integrated practice units specialising in specific disease types as described in the previous challenge means that certain procedures can move out of the hospital environment and into primary care and home care. Medical equipment that was previously used only in the hospital or clinic is finding its way into the home. For example, tremendous progress has been made since the "consumerisation" of the MEMS in developing compact, accurate, low-cost silicon sensors and actuators. This continuous innovation will support diagnostic and treatment in integrated practice units, while supporting recovery and health sustainability at home.

3.4.3.3.2. Vision and expected outcome

The transition to a healthcare carried out outside conventional hospitals requires integration of solutions and services for specific disease groups with hospital units to optimise patient-generated health data (PGHD: continuous monitoring, clinical trials at home, etc), enhanced by the integration of heterogeneous devices and systems used at home covering parts of the care cycle (smart body patches, monitoring implants, remote sensing, etc). Solutions are needed that can be integrated into secure health digital platforms, portable end-user devices, remote e-healthcare and AI front-ends.

In addition, the pharmaceutical market is experiencing strong growth in the field of biologics (genomics and proteomics, as well as microarray, cell culture, and monoclonal antibody technologies) that require preparation prior to administration. Smart drug delivery solutions are now based on innovative medical devices for the automated and safe preparation and administration of new fluidic therapies and biologic drugs. These use advanced ultra-low power microcontrollers that control the process reconstitution of the drug based on parameters identified by the practitioner, together with wireless communication modules to transmit data and ensure the patient and treatment are monitored. Smart drug delivery will improve drug adherence as patients will be empowered to administer expensive and complex drugs in their own home.

In this emerging context, care solutions need to be integrated, combining information across all phases of the continuum of care from many sources – preventing, preparing, and providing care based on person-specific characteristics. This will support the development of applicable biomedical models for specific disease groups, for customer groups and for populations, taking heterogeneous data involving history, context, or population information into account.

3.4.3.3.3. Key focus areas

Supporting prevention, diagnosis and aftercare with sensors and actuators to ensure efficient medical decision, leveraging edge computing and imaging as described in the previous section, will be crucial. The next generation of devices will incorporate increasingly powerful edge computing capabilities. Analysing PGHD from medical devices can be synchronised with a web-based monitoring system. When aggregated, this data can be then sent to the organisation's health data analytics system to process the results and compare them to previous measurements. If the analysis uncovers negative trends in the patient's health status, it will automatically notify the care team about possible health risks. The ECS industry can play an important role here in bringing ambulatory monitoring to the next level. The following enabling technology platforms can contribute to this:

- Low-power technology for sensors, microprocessors, data storage and wireless (microwave, optical, sound) communication modules, etc.
- Miniaturisation and integration technologies for sensors, microprocessors, data storage, and wireless communication modules, etc.
- Advanced sensing technologies for multiplex, painless sensing with high sensitivity and reliability.
- Printed electronics technology for textile integration and the patch-type housing of electronics.
- Low-power edge AI computing for data analysis and reduction.
- Data communication technology for interoperability of (wireless) data infrastructure hardware (wearable device connections) and software (data sharing between data warehouses for analysis, and with patient follow-up systems for feedback).
- Data security technology for interoperability between security hardware and software components (end-to-end information security).

The development of next-generation drug delivery systems will form part of the IoMT – medical devices and applications that link with healthcare systems using wireless connectivity. Smart drug delivery will improve drug adherence so that patients can administer expensive and complex (biological) drugs in their home environment. Enabling platforms are required to facilitate a transition from the legacy mechanical components seen in current autoinjectors and wearable drug delivery pumps, to highly integrated, patch-like microsystems. These include:

- High-performance sensors and actuators for drug delivery, monitoring and control.
- On-board microfluidics for in-situ preparation and delivery of formulations.
- Minimally invasive needles and electrodes for transdermal interfacing, delivery, and diagnostics.
- New materials, containers and power sources that will meet stringent environmental and clinical waste disposal standards.
- Body-worn communication technologies for IoMT integration and clinical interfacing.
- Edge AI for closed-loop control, adherence assessment and clinical trial monitoring.

The development of low-cost, silicon-based MEMS ultrasound transducers technologies is bringing ultrasound diagnostics within the reach of the ECS industry. The ECS industry has the instruments and production technologies to transform these into high-volume consumer products, something no other industry is capable of. Personal ultrasound assisted by AI data acquisition and interpretation will allow early diagnoses in consumer and semi-professional settings, as well as in rural areas. As such, they present a huge opportunity for the ECS industry. It is expected that MEMS ultrasound will enable a completely new industry, with

MEMS ultrasound transducers being the enabling platform technology that will drive things on.

Among the emerging applications of advanced MedTech, “smart wound care” – i.e. the merger of highly miniaturised electronic, optical and communications technologies with conventional wound dressing materials – will allow the treatment of chronic wounds of patients in their home without the intervention of daily nursing and/or constant monitoring of the status of the wound. While much progress has been made in wearable technologies over the past decade, new platforms must be developed and integrated to enable the rapid rollout of intelligent wound care. These include:

- Flexible and low-profile electronics, including circuits, optical components, sensors, and transducers, suitable for embedding within conventional dressings.
- Advanced manufacturing techniques for reliable integration of microelectronic technologies with foam- and polymer-based dressing materials.
- Biodegradable materials, substrates, and power sources that will meet stringent environmental and clinical waste disposal standards.
- Body-worn communications technologies for low-power transmission of wound status.
- Edge AI to assist the clinical user in data acquisition and data interpretation.

3.4.3.4. Major Challenge 4: Enhance access to personalised and participative treatments for chronic and lifestyle-related diseases

3.4.3.4.1. Status

According to the World Health Organisation (WHO) definition, chronic diseases are those of long duration and generally slow progression. Chronic diseases such as heart disease, stroke, cancer, chronic respiratory diseases, and diabetes are by far the leading cause of mortality in Europe, representing 77% of the total disease burden and 86% of all deaths. These diseases are linked by common risk factors, common underlying determinants, and common opportunities for intervention. The overall demographic change especially in Western and some Asian (Japan) countries will significantly increase the number of patients who suffer from those diseases. It will become impossible to treat all effected humans within a hospital environment, simply because of reasons of cost and required workforce which are not existent.

3.4.3.4.2. Vision and expected outcome

One of the crucial means of coping with the prevalence of the chronic diseases is to achieve a more participative and personalised approach, as such diseases require the long-term monitoring of the patient’s state, and therefore need individuals to take greater ownership of their state of health. Most chronic disease patients have special healthcare requirements and must visit their physicians or doctors more often than those with less serious conditions. Technological innovation has already been identified as a great medium to engage chronic patients in the active management of their own condition since digital health offers great convenience to such patients. Access to biomedical, environmental and lifestyle data (through cloud computing, big data and IoT, edge AI, etc) are expected to better target the delivery of healthcare and treatments to individuals, and to tailor each decision and intervention, especially for the treatment of those with multiple chronic diseases.

Patients will be connected seamlessly to their healthcare teams, care-givers and family, as treatment adherence will be more efficient with the innovations mentioned in previous sections. Remote sensing and monitoring offer great promise for the prevention and very early detection of pathological symptoms. Remote sensing and monitoring have the potential to become embedded into everyday life objects, such as furniture and TV sets, while bearing in mind the constraints related to security and privacy. Remote patient monitoring will support clinical decisions with a reduced potential for false alarms, especially for the long-term monitoring and data analysis of patients with chronic diseases.

3.4.3.4.3. Key focus areas

The ECS industry will need to take the initiative in the development of the next-generation treatment of chronic diseases. The field of remote sensing holds great promise for the lifelong and chronic monitoring of vital signs. The deployment of remote monitoring system relies on sensors integrated into bed or chair. Optical sensing techniques, for instance for remote reflective photoplethysmography, as well as capacitive and radar sensing support this approach. This will be multimodal, with fusing techniques to smart analytics to unify the data into usable information. The strength of remote sensing not only relies on the quality of the acquired signals, but also its potential to reveal slowly changing patterns – possibly symptoms from underlying physiological changes. The analysis of such datasets, currently largely unexplored, will provide new insights into normal versus pathological patterns of change over very long periods of time:

- Treatment of chronic diseases will be enhanced by an upcoming generation of small and smart implantable neuromodulator devices, which are highly miniaturised, autonomous, and cost-effective. These will be implanted, wirelessly powered by radio frequency (RF), microwave, ultrasound or energy harvesting with minimal side effects on the selected nerve through a simple and minimally invasive procedure to modulate the functions of organs in the treatment of pain management, brain disorders, epilepsy, heart arrhythmia, autoimmune diseases (immunomodulation), etc.
- Organ-on-a-chip (OOC) platforms, which lie at the junction of biology and microfabrication and biology for personalised and safer medicines, are another treatment approach, addressing, for instance, pathologies currently without effective treatment (rare diseases). Often rare diseases are chronic and life-threatening, and they affect approximately 30 million people across Europe. In an OOC, the smallest functional unit of an organ is replicated. The essential capabilities underlying the OOC field are primarily embedded microfluidics and the processing of polymers in a microfabrication environment. Smart sensors can be used as readout devices, while edge AI will be essential in data interpretation and reduction.

For chronic diseases diagnoses, LoC-based technologies – relying on miniaturisation – show promise for improving test speed, throughput, and cost-efficiency for some prominent chronic diseases: chronic respiratory diseases, diabetes, chronic kidney diseases, etc.

3.4.3.5. Major Challenge 5: Ensure more healthy life years for an ageing population

3.4.3.5.1. Status

In the last two decades, effort has been made to enhance two important and specific objectives of smart living environment for ageing well: on one hand, avoid or postpone hospitalisation by optimising patient follow-up at home, on the other hand, enable a better and faster return to their homes when hospitalisation does occur. The demographic change already introduced in the previous Challenge has two effects. Firstly, it increases the number of patients that suffer from age-related diseases, such as dementia as well as hearing & vision impairment. Secondly, it reduces the overall size of the workforce in European countries which brings the healthcare system in its current form close to a complete failure. ECS shall help to overcome this dilemma.

3.4.3.5.2. Vision and expected outcome

The following list includes some typical examples of assistance capabilities related to major chronic diseases covering the first main objective (optimisation of patient follow-up at home):

- Vital signs checker: blood pressure meter, pulse meter, oximeter, thermometer, weight scale expanding to sensor hubs for vital signs enabling applications like Electroencephalography (EEG), Electromyography (EMG), Electrooculography (EOG), Electrocardiogram (ECG), Galvanic skin response (GSR), Blood volume pulse (BVP),
- proactive speech centric HMI approaches to enable interactive operation of healthcare related devices.
- Hospital's software interface, the patient's file, the patient's risk alarm centre with automatic call to healthcare practitioners.
- Audio and Video communication support (between the patient and their nurse, doctor, and family), and interactive modules for the patient (administrative, activities, menus, medical bot chat, etc.).
- Authentication and geolocation of patients, with patient or patient's family consent.
- Teleconsultation for nights and weekends at the foot of the bed of patient's hospital or retirement home.

The second objective – smooth home return – relates to suitable technical assistance in addition to human assistance:

- Enhance the patient's quality of life and autonomy.
- Improve the patient's safety and follow-up in their room through a reinforced work organisation.
- Allow monitoring of the patient's progress to motivate them during their rehabilitation period and encourage and enable active participation in social life
- Minimally invasive therapies allowing for shorter hospital stays and improved patient wellbeing.

- First-time-right precision diagnoses to prevent hospital readmissions.
- Prepare for the return home: patient support in appropriating technical aids by integrating these solutions into rehabilitation.

Efforts are being made to enhance medical and social care services through different kinds of robots. The purpose here is to provide advanced assisted living services via a general-purpose robot as an autonomous interaction device that can access all available knowledge and cooperate with digital appliances in the home. In this sense, autonomous mobile robots offer several advantages compared to the current (stationary) Ambient Assisted Living (AAL) solutions. Due to sensor-augmented user interfaces, like speech recognition and synthesis, human computer interaction is becoming increasingly natural. Consequently, robots will come to represent a familiar metaphor for most people.

Neurorehabilitation is sometimes required after hospitalisation and is generally a very complex and challenging undertaking resulting in both “successes” and “failures” (setbacks). Neurological patients typically report having “good days and bad days”, which affect performance, motivation and stamina, and where cognitive stimulation (AI-based speech producing programs, social robots, etc), for example, has the potential to improve the efficiency of neurorehabilitation and relieve some of the pressure on health systems. Robotics is well suited for precise, repetitive labour, and its application in neurorehabilitation has been very successful. This is one of the main reasons why the rehabilitation robotics market has tripled over the last five years and, today, rehabilitation robotics is one of the fastest growing segments of the robotics industry. This industry is dominated by European companies that can deliver highly innovative solutions with a strong scientific basis and exceptional manufacturing quality. Based on market size and need, it is projected that the compound annual growth rate (CAGR) for rehabilitation robotics will soon reach between 20% and 50%.

3.4.3.5.3. Key focus areas

The ECS industry can significantly upscale the “ageing well” area, as it is enabled by most of the focus topics developed in the previous sections. The industry is playing an important role in bringing ambulatory monitoring to the next level. Important aspects here are reducing costs, improving user friendliness (e.g. easy to wear/use devices, interoperable gateways, reduction of patient follow-up systems) and data security.

The enabling technology platforms detailed below are expected to significantly contribute to this prevalence of the ECS industry in ageing well, taking into account that ageing well is very much related to “ageing in place”:

- Low-power technology for sensors and actuators, microprocessors, data storage and wireless communication modules, etc.
- Miniaturisation technology for sensors and actuators, microprocessors, data storage and wireless communication modules, etc.
- Printed electronics technology for textile integration and patch-type housing of electronics.
- Low-power edge AI computing for data analysis and data reduction.

- Data communication technology for interoperability of (wireless) data infrastructure hardware (wearable device connections) and software (data sharing between data warehouses for analysis and with patient follow-up systems for feedback).
- Data security technology for interoperability between security hardware and software components.
- Robotics systems enabling patients to overcome loneliness or mental healthcare issues, complemented with cognitive stimulation by AI-based speech recognition and synthesis enabling augmented HMI for social robots

Interoperability is surely the main challenge faced by the ECS industry in achieving full impact due to the vast heterogeneity of IoT systems and elements at all levels. Interoperability and standardisation need to be elaborated in relation to data and aggregated information. Thus, it is not enough to be able to receive a message, i.e. to understand the syntax of the message, but it is also necessary to understand the semantics. This requirement implies the development of a data model that maps semantic content from the data received from devices into an information system that is usually utilised for collecting and evaluating data from monitored persons. It must be based on several relatively simple principles: creation of formats and protocols for exchange of data records between healthcare information systems; format standardisation and connected interface unification; improvement of communication efficiency; a guide for dialogue between involved parties at interface specification; minimisation of different interfaces; and minimisation of expenses for interface implementation.

3.4.4. TIMELINE

The following table illustrates the roadmaps for Health and Wellbeing.

MAJOR CHALLENGE	TOPIC	SHORT TERM (2025–2029)	MEDIUM TERM (2030–2034)	LONG TERM (2035 and beyond)
Major Challenge 1: Enable digital health platforms based upon P4 healthcare	Topic 1: establish Europe as a global leader in persona-lised medicine deployment	• IoMT-enabling patient-generated health data • Expansion of AI on the edge • High level of digital trust – privacy and security by design, hardened embedded AI models	• Development of multimodal data analysis • Improvement of energy efficiency (energy harvesting, etc.) • Secure digital health platforms, portable end-user devices, remote e-healthcare, and AI front-ends	• New tools for clinical decision-making and precision medicine • Scalable digital health platforms

Major Challenge 2: lead the healthcare system paradigm shift from treatment to health promotion and prevention	Topic 2: enable the shift to value-based healthcare	• Disease detection from biomarkers derived from medical images and sensors • Predictable and repeatable outcome of diagnostic imaging • Digital supply chains, automation, robotics, and next- generation interoperability • Early diagnosis through PoC diagnostic systems	• Clinical decision-making augmented by a combination of electronic medical records[1], imaging, biomedical models • EHRs supporting precise communication between different care-givers, PoC diagnostic systems or AI-based clinical decisions • Efficient healthcare information infrastructure, lowering costs while improving health outcomes	• Shift from general hospital to specialised integrated practice units • Next generation of smart minimally invasive devices • Disease detection from biomarkers derived from medical images • Outcomes cover the full cycle of care for the condition, and track the patient's health status after care is completed
Major Challenge 3: home becomes the central location of the "healthcare consumer"	Topic 3: build an integrated care delivery system	• Use heterogeneous data from more sources (patient-generated health data, edge computing, and imaging to ensure efficient medical decisions, etc.) • Remote decentralised clinical trials development (smart body patches, monitoring implants for continuous monitoring, etc.)	• Next-generation drug delivery systems (highly integrated, patch-like microsystems) will form part of the IoMT	• Care solutions integrated, combining information across all phases of the continuum of care, preventing, preparing, and providing care based on person-specific characteristics • Holistic healthcare involving all imbalanced health situations of the patient

[1] Electronic medical record (EMR): A computerised database that typically includes demographic, past medical and surgical, preventive, laboratory and radiographic, and drug information about a patient. It is the repository for active notations about a patient's health. Most EMRs also contain billing and insurance information, and other accounting tools.

Major Challenge 4: ECS industry supports EU strategy to tackle chronic diseases	Topic 4: enhance access to personalised and participative treatments for chronic and lifestyle-related diseases	• Accurate long-term monitoring and data analysis of patients with chronic diseases and co-morbidities • Make treatment adherence more efficient (smart drug delivery based on innovative medical devices,etc.)	• Development of active or passive implantable medical devices	• OOC platforms addressing pathologies currently without for chronic disorders effective treatment (rare diseases)

Major Challenge 5: ECS industry fosters innovation and digital transformation in active and healthy ageing	Topic 5: ensure more healthy life years for an ageing population	• Expand the focus of Medicare beyond treatment of diseases towards prevention of diseases • Optimization of consumer/patient follow-up at home to support ageing in place (remote patient monitoring, geolocalisation, etc.), • enable the silver generation to maintain their active participation in social life by mitigating the effects of dementia, hearing and vision impairment and avoid stigmatisation and isolation	• P4 healthcare – predictive, preventive, personalised, participatory – as opposed to reactive healthcare • Suitable technical assistance in addition to human assistance (humanoid robots, advanced assisted living, rehabilitation robotics, etc) • Precision diagnosis to prevent hospital readmissions, • Enable personal assistants through audio, visual/ gesture-based HMI to interact with consumers/ patients/disabled person to overcome physical impairments /handicaps limiting the self-determined management of daily life.	• Changing the way "Health Consumers" and "Patients" expect to interact with healthcare. • Data model diffusion that maps semantic content from the data received from devices into an information system that is usually utilised for collecting and evaluating data from monitored persons. • enable devices supporting HMI, which leverage neurological and biometric data like Electroencephalography (EEG), Electromyography (EMG), Electrooculography (EOG), Galvanic Skin Response (GSR) and Blood Volume Pulse (BVP) to evaluate brain waves, facial expression and jaw & eye movements, emotional arousal and heart rate and heart rate variability.

3.5

ECS Key Application Areas

AGRIFOOD AND NATURAL RESOURCES

3.5 Agrifood and Natural Resources

3.5.1 Scope

In recent years , the conditions of the planet have changed abruptly as predicted by the Intergovernmental Panel on Climate Change (IPCC). According to the latest report from IPCC[1], global warming is causing major changes in precipitation patterns, oceans and winds, in all regions of the world and, in some cases, irreversibly so. The intensification of natural phenomena is putting the viability of life on Earth at risk, and from now on will have serious repercussions on food security, health, and sustainable development. Extreme hot temperatures in normally cold countries; melting of the poles at an accelerated rate, and consequently, rising sea levels threatening coastal areas; prolonged droughts in previously fertile and productive places on different continents; scarcity of fresh and affordable water for human consumption in large cities are just some of the issues we are now regularly facing. Moreover, forest fires have doubled worldwide in the last 20 years destroying around 3 million hectares each year particularly in boreal forests, leaving the largest climate change related carbon deposits on the planet so far according to a study conducted jointly by three institutions: Global Forest Watch (GFW), World Resources Institute (WRI) and the University of Maryland (UMD). The study concludes that we will lose or degrade these important lungs in the medium term despite these forests being one of our best defences against climate change and agriculture with agroforestry practice.

There is a strict relationship between climate change and agriculture. The two-way relationship of climate change and agriculture is of great significance because we need to adopt effective practices to mitigate risks to human health and crop production. Because of these reasons, the term Climate-Smart Agriculture[2]has recently been adopted to describe the innovative use of technologies in agriculture as highlighting the fact that it is important to have smart/precise agriculture, but it has to be respectful of climate. The actual climate-agriculture interaction can be considered a lose-lose exchange, being climate problematic for agriculture and because agriculture is one of the main reasons for greenhouse gas emissions. New smart technologies and increased agroforestry practices can change this relationship to positively impact both sides.

Moreover, two other emergent processes are becoming relevant: desertification and tropicalization[3]. The last one is related to multiple climate-induced range shifts, including the expansion of tropical species towards the poles and concomitant loss of temperate species from warming areas. This alters the species interactions and may lead to changes in marine biogeographic structure, with a lot of consequences for general ecosystem functioning and fisheries system.

On one hand, as explained above, the impact of climate change on agriculture is mainly related to extreme heat events,reduction in precipitation and/or intensive flooding and water availability, that result in

---

[1] IPCC Report: Climate Change 2023 https://www.ipcc.ch/report/ar6/syr/

[2] https://www.fao.org/climate-smart-agriculture/en/

[3] The ecological and evolutionary consequences of tropicalisation, Trends in Ecology & Evolution, March 2024, Vol. 39, No. 3

decreased crop productivity. On the other hand, the impact of agricultural activities to climate change is related to two specific factors:

- Farming in particular releases significant amounts of greenhouse-gas emissions, in particular methane and nitrous oxide. The agriculture sector alone represents almost a quarter of global emissions.
- Agrochemicals released to fight against pests contaminate soils and waters as a direct consequence of the use of these substances.

However, climate change is only one of the many problems that agriculture must face[4]. In fact, growing global demand and competition for resources, food production and consumption need to be redesigned in a proper way, linking agriculture, energy, and food security.

Consequently, contemporary economic and ecological challenges mean our food production must support a new balance between production in quantity and production of quality. Achievement of this new balance in food production exposes us to risks in various forms (war, market fluctuations, large scale public health and animal health) . These risks must be addressed to obtain a transition towards a more sustainable and inclusive food system from farm to fork. This will require significant actions such as reducing food loss and waste, adopting dietary changes, and adapting how we use arable land. These actions will help industry meet global food needs while safeguarding farmers' livelihoods as well as contributing to decarbonisation and climate change stabilization.

As a primary sector, digitalization of agriculture is not trivial because of the great variability (e.g. climate or other natural phenomena, farm typology) of events on crops and land. In particular, the diversity of farming systems and farmers is essential for targeting agricultural interventions in any mixed crop-livestock farming system. . Moreover, farmers will need to change their behaviour. Service support and maintenance with adequate education on new technologies must be introduced, as well as a robust and reliable precision farming infrastructure.

The G20 Ministers of Agriculture, assembled on 16-17 June 2023 in Hyderabad, India[5], emphasized their commitment to food security and nutrition for all, through the development of inclusive, resilient, and sustainable agriculture and food systems, with the need to work together to promote food security and nutrition. The G20 meeting recognized that the current crises (lastly the war in Ukraine) are multi-dimensional and therefore, require a multi-layered approach, combining coherent and effective short-, medium-, and long-term responses in the spirit of "One Earth, One Family, One Future", tackling all crises with the same urgency. It is expected that this message will be repeated in the G20 meeting in November 2024 which is after the publication date of this document.

Among the high-level principles established by G20, Principle 6 is related to the need of Acceleration of Innovation and the Use of Digital Technology in Agriculture. This is relevant because it is a further emphasis to promotion of development and safe application of digital tools tailored to the various needs

---

[4] Agriculture and Climate change, European Environment Agency, 30 June 2015 https://www.eea.europa.eu/signals/signals-2015/articles/agriculture-and-climate-change

[5] https://g7g20-documents.org/database/document/2023-g20-india-sherpa-track-agricultural-ministers-ministers-language-g20-agriculture-ministers-meeting-outcome-document-and-chairs-summary

of the agriculture sector. The importance of strengthening digital solutions to empower all farming communities, including smallholders, was recognized.

All of this creates many opportunities for the ECS community to contribute to the disruptive move of the agrifood sector towards a sustainable future. Innovations for, and digitalization of agriculture should be available as soon as possible to bring a new level of agri-food system resilience, capable of having a more productive, decarbonised, and sustainable agriculture globally.

Smart Internet of Things (IoT) systems have become very important for sustainable production and consumption of safe and healthy food, as well as for sustainable practices in agriculture, livestock, aquaculture, fisheries and forestry. They can foster access to clean water, fertile soil and healthy air for all, in addition to helping fighting against pests while preserving biodiversity and restoring the planet's ecosystems. In short, the use of these connected objects (IoT) helps the stakeholder to increase productivity while ensuring sustainability. Finally, IoT systems should provide innovative GHG emissions tracing solutions to facilitate decarbonization. The rise of AI will also allow for novel digital support systems, like such as expert advice systems built on Large Language Models like ChatGPT which not only refer to knowledge bases but also to live data provided by afore mentioned IoT systems.

Digital farming support should also considered novel development like the digital product pass which supports circular economy in other domains but may also be an important tool to support related topics like the variable due date.

In this Chapter, five Major Challenges have been identified. The first two Major Challenges relate to livestock and crop health, connected to farming systems and food supply chain assurance and management. For instance, IoT system technologies can be used in pest management or towards minimising the use of pesticides and antibiotics. Farming systems and food supply chain management benefit from smart IoT systems, including the use of traceability frameworks with trustworthy and security features[6], as well as from robots and drones, to revolutionise modern agriculture and food production. The third Major Challenge addresses issues such as soil health, air quality and the environment, all in terms of smart integrated monitoring technologies, and the use of smart waste management systems and remediation methodologies. The objective is to protect the environment to reduce the destruction of ecosystems caused by a myriad of anthropogenic activities and to reduce GHG emissions. A large decrease in the use of chemical fertilizers is a major aspect. These fertilizers are made artificially from soil-essential macronutrients like nitrogen, phosphorous, and potassium and they may contain ammonium sulfate, urea, potash and ammonia, among other substances, depending on their structure and the crops and soils for which they are intended.

Great advantages of these fertilizers are undeniable related to high production per hectare, they can be a boost to the health and expectations of plants in advanced stages of cultivation. Nevertheless, main drawbacks are linked to soil degradation, groundwater contamination and salt burns. Therefore, it's relevant to take into account among the challenges the way to support the reduction of these chemicals.

---

[6] K. Demestichas, N. Peppes and T. Alexakis. "Survey on security threats in agricultural IoT and smart farming." Sensors 20.22 (2020): 6458

The reduction of fertiliser dependency is also encouraged by the EU Member States' Common Agricultural Policy (CAP) Strategic Plans. The fourth Major Challenge refers to the key role that IoT systems can play in water quality monitoring and management and access to clean water. An important aspect here is the overall management of water usage, as well as smart treatments to foster the circular use of wastewater, rainwater and storms/floods.

The fifth Major Challenge relates to biodiversity restoration for ecosystem resilience, how electronic components and systems (ECS) can contribute to the restoration/preservation of a greater variety of crops, and greater fauna and flora species diversity, to ensure the natural sustainability of healthy ecosystems (agriculture, aquaculture, fisheries and forestry) by enabling them to better withstand and recover from misuse, abuse or disasters.

All five Major Challenges in this Chapter align with key Horizon Europe missions, as well as the European Green Deal and Digital Europe. To efficiently address these challenges, significant advances are crucial in the fields of new materials, manufacturing technologies, innovative sensing solutions, information and communications technology (ICT), Artificial Intelligence (AI), robotics, energy management, harvesting and transfer, electronics and photonics, and other technologies, as well as in circular industries. These challenges also address most of the technologies required to support the decarbonization actions in the farm proposed by McKinsey[7] to achieve the IPCC 1.5° C pathway[8]. Together with the best management of agricultural resources and the good practises for a sustainable agriculture, all the challenges form a great occasion for farmers to change direction to minimize damage to the environment, while taking into account the economic needs of the farmers themselves.

Figure 3.5.1 illustrates the main challenges that our society is faced with a) the demand shift from resource-intensive consumption to resource-efficient consumption and b) markets shift from low-connectivity to high-connectivity solutions. Both are required to reach an open-source sustainability.

[7] The agricultural transition: Building a sustainable future, McKinsey & Company, June 2023

[8] IPCC Sepcail report "Global Warming of 1.5 °C pathway, Korea 1-5, October 2018

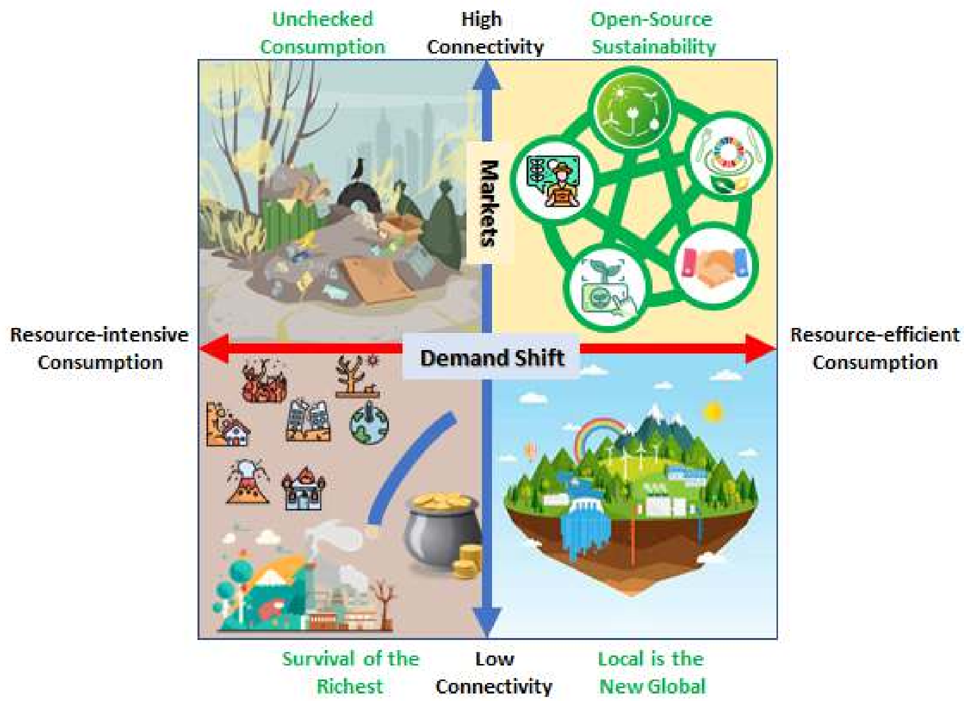


Figure 3.5.1 - The Scenarios: Four Potential Future Worlds[9]

3.5.2 Application trends and societal benefits

External requirements

According to the UN[10], if the global population reaches an expected 9.8 billion by 2050, the equivalent of almost three Earth planets could be required to provide the natural resources needed to sustain current lifestyles. Increasing food production is driven not only by population growth, but also by more demanding and sophisticated diets, with zero net emissions, as populations become wealthier. On the other hand, productivity is being hit hard by climate change in regions where food scarcity and inefficient resource management is most prevalent. The necessary acceleration in productivity growth is being hampered by the degradation of natural resources (including soil), a reduction in biodiversity, and the spread of transboundary pests and diseases of plants and animals, some of which are already becoming resistant to antimicrobials[11]. Investments in changing agricultural practices and incorporating technological innovation has boosted productivity, but the yield growth is far from sufficient. A more holistic and innovative approach is needed to reduce the strain on natural resources and enhance their quality, while also increasing food productivity. At the same time, food losses and waste claim a significant proportion

---

[9] Shaping the Future of Global Food Systems: A Scenarios Analysis, World Economic Forum White Paper

[10] Microsoft Word - Key Findings WPP 2017_Final EMBARGOED (un.org)

[11] 2017 FAO. 2017. The future of food and agriculture – Trends and challenges. Rome

of agricultural output, whereas poor bio-waste management and packaging increases environmental pollution.

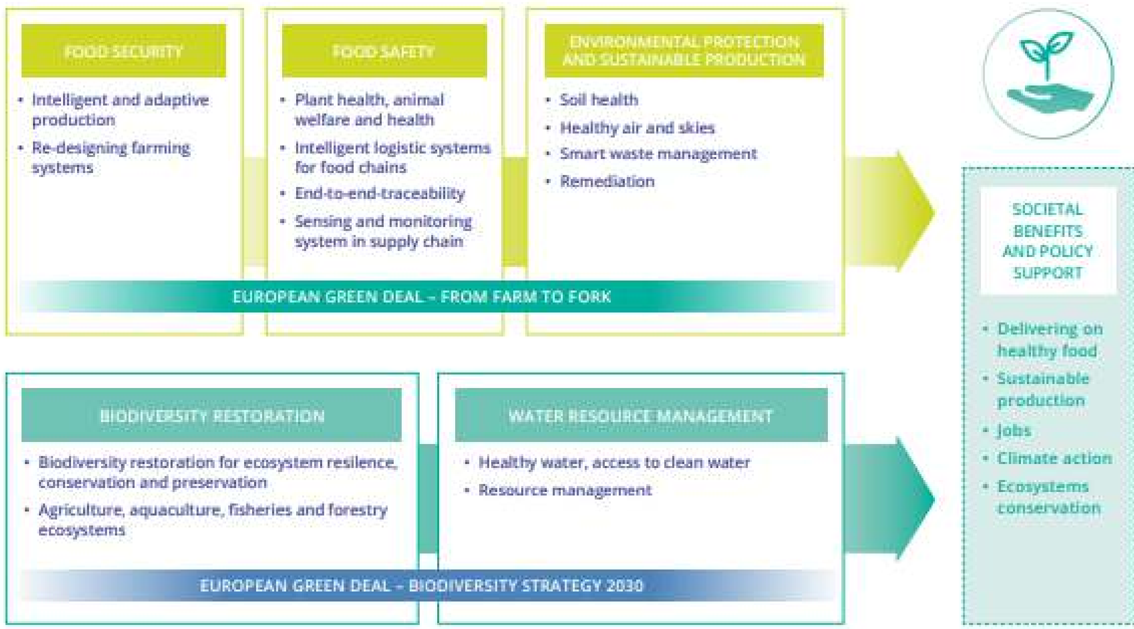


Figure 3.5.2 - Main Agrifood and Natural Resources goals and associated challenges

Addressing key issues on food security and sustainable production would lessen the need for production increases while improving the natural resource base. For instance, reducing GHG emissions is a priority because they significantly contribute to climate change. Three major sources (land-use change, enteric fermentation and energy use) combined account for almost 74% of the total GHG emissions. Other examples are, mitigating the effect of natural and human pressures on water bodies, namely by reducing general pollution and plastics, eutrophication, acidification and warming-up as much as possible. Less than 2.5% of the world's water is fresh[12], and water pollution in rivers and lakes is occurring faster than nature can recycle and purify. Currently, more than 2 billion people live with the risk of reduced access to freshwater resources[13], and by 2050 at least one in four people is likely to live in a country affected by chronic or recurring shortages of freshwater. Now, 2.6 billion people are economically dependent on agriculture[14] despite 52% of arable land being moderately or severely affected by soil degradation. Air quality has also been deteriorating in both rural and urban areas because of the spread of particulate matter in addition to the release of greenhouse gases (GHGs) all of which have detrimental effects on the population[15] and on the climate.

---

[12] All about water, https://www.iaea.org/sites/default/files/publications/magazines/bulletin/bull53-1/53105911720.pdf
[13] UN 2019 The Sustainable Development Goals Report 2019: Goal 6: Clean water and sanitation / https://unstats.un.org/sdgs/report/2019/The-Sustainable-Development-Goals-Report-2019.pdf
[14] ibid: Goal 15
[15] https://www.eea.europa.eu/themes/air/health-impacts-of-air-pollution

Today, farmers still spread much more fertiliser than is required on their fields. Consequently, excess nutrients such as nitrates and phosphates accumulate in the soil and filter into groundwater with a dramatic impact on the environment and public health. Therefore, there is increasing pressure on the agricultural sector to find decarbonized and sustainable solutions for reducing environmental pollution caused by fertilisation, pesticides, livestock and energy production emissions. Smart production processes and intelligent logistic systems across the whole supply chain are some of the solutions that can yield optimisations to reduce emissions with increased productivity, while ensuring safe food production. This is particularly important given that currently, every year, almost one in 10 people fall ill due to food-borne diseases.

The pandemic crisis has shown the vulnerability of the overall agri-food supply chain when compromised by employee illness or travel restrictions enforced by the lockdown constraints. These circumstances appear unprecedented, but they are relevant for every type of pandemic that could occur worldwide, crossing international borders and affecting large populations.

New epidemiological methods that utilize dynamic network analysis[16] to analyse the key drivers of emerging pathogen movement are needed. In the agrifood sector, this is compulsory, because the overall agrifood elements are strictly interconnected (from seeds and plants to livestock management and crop production, as well as postharvest transportation are single point of risks of pathogen and mycotoxin movement in stored food).

Effective surveillance strategies are essential to support agrifood health programs, crisis prevention, improvement in biosecurity and agriculture decarbonization. Here also, ECS solutions are the efficient way to contribute to all these topics.

Societal Benefits

In response to an ever-increasing set of challenges faced by the world, the UN defined 17 Sustainable Development Goals (SDG) to act as a blueprint for achieving a better and more sustainable future for all. The SDG implementation plans (SDG 2, 6, 12 and 15 are particularly relevant) to address the global challenges we face in protecting biodiversity, our natural resources and acting on climate change. Furthermore, it includes actions relating to socioeconomic drivers aimed at eliminating poverty, hunger, inequality, and achieving responsible consumption and production, sustainable prosperity, peace and justice. In Europe, national and EU policies such as the **"From Farm to Fork"**[17] and **"Biodiversity Strategy 2030"**[18]**,** reflect and amplify the underlying SDG objectives with a set of measures – from regulatory frameworks to incentives and investments for development, and the deployment of holistic innovative approaches in a circular economy, agroecology, agroforestry, climate- smart and sustainable agriculture, bioeconomy, and the Blue Economy.

[16] The persistent threat of emerging plant disease pandemics to global food security at https://doi.org/10.1073/pnas.2022239118

[17] https://ec.europa.eu/food/sites/food/files/safety/docs/f2f_action-plan_2020_strategy-info_en.pdf

[18] https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020DC0380&from=EN

3.5.3 Major Challenges

This section discusses the five Major Challenges that need to be addressed in the domain of agriculture (food security, food safety, environmental protection and sustainable production), natural resources and biodiversity, and how smart IoT systems and associated key enabling technologies can help achieve them.

- **Major Challenge 1**: Food Security.
- **Major Challenge 2**: Food Safety.
- **Major Challenge 3**: Environmental protection and sustainable production.
- **Major Challenge 4**: Water resource management.
- **Major Challenge 5**: Biodiversity restoration for ecosystems resilience, conservation, and preservation.

3.5.3.1 Major Challenge 1: Food Security

Food security[19] and food safety[20] are two complementary/interdependent concepts that are characterized in different ways: one indicates the economic and social security of availability of food supplies; the other indicates their health and hygiene safety. Together, these terms refer to the same processes linked to agricultural production, and it is relevant to manage them simultaneously.

Figure 3.5.3 presents the interrelation between food security and foot safety concepts, as well as their main constituent elements. This section and the next will address the challenges related to food security and food safety from an ECS perspective.

Figure 3.5.3 - Food security and food safety

---

[19] Food security has been defined by the FAO as "Food security exists when all people, at all times, have physical and economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life"

[20] Food safety is an umbrella term that encompasses many facets of handling, preparation, and storage of food to prevent illness and injury. Included under the umbrella are chemical, microphysical, and microbiological aspects of food safety

3.5.3.1.1 Status, vision and selected outcome

Consolidated advances in Industrial Internet of Things (IIoT) have already started to shape smart manufacturing in the food and beverage[21] industry. Access to relevant and role-based information, in real-time or near real-time, is key to ensuring the efficient storage and processing of data, and their appropriate use for optimised decision-making at every level of next-generation automation systems and robotics, e.g. cyber-physical systems (CPS). Therefore, sustainable production, safety and quality do not only depend on the product itself, but they also depend on respective processes and their control as offered by key data gathering and monitoring, smart autonomous sensing, data analysis,diagnostics and predictive maintenance, and control systems. Ultimately, intelligent food production frameworks can consider consumer needs in specific markets, and such systems can provide intelligent recommendations for adjusting the amount and quality of food, accordingly, assuring food security (i.e. enough food for each market, avoiding food loss) and food safety (i.e., healthy food), while also considering environmental concerns and societal impact, paying attention to the food traceability process as well, because it is a real core of any quality assurance for agrifood sector.

In particular, smart sensing had a recent important benefit thanks to the appearance of low-cost, low-power, self-powered solutions, giving the possibility of not only remotely monitoring the crops (satellites, multi-spectral cameras), but to install the sensors inside the crops and plants, for receiving directly from the site the data needed to understand the health status of plants and soil. A LinkedIn post from World Economic Forum indicated the technology of Plant Wearable Sensors as one of the 5 technologies about to change the world[22], and as one of the key solutions to increase food production by 70% by 2050 to be able to feed the world population.

Following the trend in manufacturing industries, digital twins[23] are the next step for the food industry and farming systems. In short, digital twins allow digital/virtual representations (models) of physical objects and processes, coupled with behaviour models that enable simulation and prediction upon changes to variables associated with the objects or the surrounding conditions. Digital twins are remotely and real-time connected to the objects in the physical world to reflect the dynamics of real systems. Thus, digital twins are expected to take the farming and food industry to the next level in terms of productivity and sustainability. As a use case example in precision farming, a digital twin can be used in the event of a plague infection for simulating the effect of applying multiple alternatives, taking into account the current condition of the crops, the available biological models, the expected evolution of weather conditions, etc. to figure out what is the optimal treatment (and timely application) against the plague in order to minimize both the impact in productivity and the environment footprint of the treatment. Essentially, digital twins will become ultimate decision-making optimisation tools by integrating production process variables and market and consumers variables thereby avoiding food loss. In any case, development of digital twins in farming is far behind its counterpart in the manufacturing industry for several reasons

---

[21] Beverage will be considered as food in the rest of the document.

[22] https://www.linkedin.com/posts/world-economic-forum_wef24-ugcPost-7149142642344296448-lxGR/?utm_source=share&utm_medium=member_ios

[23] https://www.sciencedirect.com/science/article/pii/S0308521X20309070

because physical objects in farming are “living” objects (crops, trees, animals…), and because variables of interest in farming are highly heterogeneous and are complex to model and measure.

3.5.3.1.2 Intelligent and adaptive food production

To develop intelligent food production systems, solutions are required in (but not limited to) the following fields:

- In-line inspection, networked packaging systems and robot technology in the warehouse to allow for a smart workflow to manage, monitor, optimise and automate all processes accordingly.
- Intelligent control room systems to enable correlations between machine malfunctions and load parameters to be detected immediately, thereby enabling maintenance work to be carried out early and on schedule, with a reduction in costly downtimes.
- Food industry imposes specific requirements (e.g. in food processing) that may take advantage of smart bio-sensing high-quality monitoring to reduce the amount of water and chemicals used in such processes, and to prevent contamination.
- AI/machine learning (ML) and big data models must be devised and used to offer further intelligent decision-making and, whenever possible, should be employed directly at-the-edge for greater energy efficiency.
- IIoT systems, based on AI and digital twin technology, can provide the flexibility to tailor-make new products to help cope with ever demanding diets.
- Integration or combination of IIoT systems with Large Language Models to generate a new generation of smart expert advice systems

3.5.3.1.3 Re-designing Farming Systems

Precision Farming Systems

Advanced farming machines and robotic collaborative systems are needed for cost-effective land and livestock management, as well as for large-scale arable and fruit crops management. Tasks can be performed in parallel, enabling economies of scale. Advanced machines include the following:

- *Robotic systems (for harvesting, weeding, pest control, pruning)*: autonomous robots or swarms of light(er) robots (causing less soil compaction) can replace intensive and strenuous labour practices as the worldwide population transitions from rural to urban areas and manual labour declines as farmers get older. These low cost, adaptive, edge-AI based, multifunctional robots provide high levels of labour input to deliver productive, biodiverse, heterogenous landscapes of diverse cropping systems. Agricultural robots need to be equipped with improved capabilities for sensing and perception. This aspect should be accelerated to tackle this problem and increase efficiency. Special attention must be paid to safety and trustworthiness aspects for those robots expected to work collaboratively with humans or close to livestock.
- *Drones:* remotely piloted or autonomous unmanned aerial vehicles (UAVs), either flying alone or in swarms, can mainly improve efficiency in two application areas: (i) monitoring large areas with intelligent computer vision devices to provide a higher level of detail and on-demand

images, especially as drones can overcome limitations of satellite imagery (e.g. images below forest cover); and (ii) in the use of phytosanitary (plant health) products to increase efficiency and reduce environmental impact by avoiding indiscriminate chemical dispersion and following predetermined prescription maps.

- *Satellites:* these allow for improved information regarding fields, although a combination of data from further sensors with increased update frequency, improved performance and spatial resolution would also be needed. Moreover, small satellites (micro, nano) could provide larger IoT connectivity services when internet coverage may not be available at any location where a WSN network is to be deployed. Hence, low power and low-cost solutions are needed where either links can be established or where data can be collected for later processing in powerful backends.
- *Wireless sensor networks (WSNs)* and smart actuators deployed across fields will form the backbone of heterogeneous - multi-agent - collaborative approaches. Local parameters (e.g. ambient temperature, soil pH, soil salinity, relative humidity, etc.) measured from multiple sensors planted in the soil or attached to the plants could be retrieved remotely by e.g. drones and/or robotic systems to deepen the field analysis provided by image-based techniques. Coupled to the proper AI and decision systems, WSNs will also further help in automatically triggering the appropriate actions (e.g. drones could locally release agrochemicals after interrogating/analysing sensors, water irrigation systems could be activated only in some land areas, etc.).
- *Decarbonization:* To achieve zero net emissions in the different agriculture areas carbon verification and monitoring tools are required to measure carbon emissions and sequestration. Decarbonization is also requiring robust and trustable measurement methodologies, allowing producers, but also policymakers, on one side to control the quantity of carbon emitted and on the other side of carbon sequestrated, allowing trustable information about decarbonization control. As stated in the FAO document "Carbon sequestration in dryland soils ", "... continuous monitoring of carbon losses and gains in the farming system must be an integral part of a project for which a designated local institution could be responsible".
- *Digital farming Support as a Service (FAAS):* Most of the farms in Europe are small scale, below 10 hectares, whilst only about 1% of all farms are above 500 hectares[24]. An important challenge remains that all sizes of farms, including small and medium sized, should have access to digital solutions, namely cost-effective ones, and to facilities to easily exploit them. Whilst large scale farms have the means to setup and maintain large infrastructures and even robotic appliances, smart scale farms should benefit from the provision of digital support solutions, for instance, farm-monitoring via sensors, local maintenance, virtual cooperation, precision agriculture applications, etc. - as a service. This kind of services should be created and provided through local cooperatives or new service providers specialised in the HW infrastructure and SW applications involving several types of expertise in agronomy, communication, data analytics, computer science, etc. Having digital farming as a service, business models may help to finance CAPEX while technology transition to provide better cost model scenarios and facilitate the deployment of this

---

[24] Destatis Statistisches Bundesamt, Betriebsgrößenstruktur landwirtschaftlicher Betriebe nach Bundesländern, retrieved October, 6, 2022 from https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Landwirtschaft-Forstwirtschaft-Fischerei/Landwirtschaftliche-Betriebe/Tabellen/betriebsgroessenstruktur-landwirtschaftliche-betriebe.html

technology in the European farms. As with most non-ECS domain experts, famers would benefit by digitalization expertise being provided by ECS tools and services supporting e.g. experience sharing, education and remote assistance. This can include digital ecosystems, remote support via XR technology, or simple interactive support systems using technologies like LLM, e.g. ChatGPT, to facilitate existing knowledge which can be effectively queried

Couplings[25] between the technologies cited above and the data sources are possible and make it possible to enrich knowledge and respond to other field issues. Take the example of the drones above. Georeferenced imagery data by drones can be completed:

- on wider perimeters but with lower frequencies by satellite imagery.
- by images acquired by proximity sensors onboard self-propelled vehicles (tractors or robots) and their associated equipment during interventions in the plots as well as during pedestrian observation phases by farmers.
- by data collected via networks of communicating sensors covered by the Connectivity Chapter.

All the proposed solutions should meet important requirements such as cost-efficiency, compactness, reliability, lifetime, low power, security, interoperability with existing machinery and between systems implementing appropriate security schemes and taking human factors into account. Furthermore, training systems based on virtual, augmented, and mixed reality and simulators are needed for training people (e.g. operators), independent of seasonality or safety issues.

Horticulture/greenhouses, urban and vertical indoor agriculture, and agrovoltaics

Urban agriculture is being promoted as a promising option for sustainable food, a better quality of life, and community engagement. The goal of this modern version of agriculture is to grow and deliver high-quality food with a minimal waste of resources.

Many crops in vertical indoor farms are often cultivated using hydroponics, a technique where there is no need for soil and fertiliser as the growing plants are supplied with irrigation water. In fact, recent environmental challenges have promoted the intensification of “soil-less agriculture” in an urban context to decrease the negative impact on nature. Even if hydroponics produces quality crops with high efficiency, there is an area of opportunity here to better monitor and control the fertiliser components in the irrigation water, such as through the development of:

- novel and low-cost online sensors for optimised control, such as nutrient sensors to enable smaller discharge of fertiliser into natural waters.
- robots with a high precision level to perform automatic harvesting to reduce the overall production costs, which are currently high, to be competitive with traditional agriculture.

---

[25] See in French https://blog.irt-systemx.fr/ameliorer-la-production-du-secteur-agricole-avec-les-nouvelles-solutions-technologiques/

- autonomous indoor farming systems in which cultivation is controlled remotely via AI, based on measurements of crop properties with the help of intelligent sensors and (edge) AI-based digital twin models of such plants.

Climate change is causing high temperatures, torrential rains and hail, causing significant damage in large crops. The introduction of **Agrovoltaics** has the potential to sustainably increase agricultural yields, reduce water use, create additional revenue, and promote equity for small-scale farmers[26]. Additionally, solar panels can provide energy directly to farms, reducing their dependency on fossil fuels and encouraging energy independence for small-scale farmers in developing communities; excess energy can be sold to the grid. The shade provided by the panels and the possibility to add environment sensors and ad hoc irrigation systems can make farms more water-efficient and provide valuable shade for livestock, leading to greater productivity for both crop and animal yields.

3.5.3.2 Major Challenge 2: Food Safety

3.5.3.2.1 Status, vision and expected outcome

Application of high-tech sensors and AI to monitor, quantify and understand individual plants and animals, as well as their variability, to ensure food safety is key for next generation novel ecology-based agricultural systems. Smart sensors and monitoring technology that can adapt to the unpredictability and variation of living systems are required. This Major Challenge will require integrated digital technology solutions such as ecology-based robotic systems that can control the bio-physical processes (including growing conditions) and understand the biological environment (for plants and animals). However, innovative ecology-based robotic systems' manipulation of operations is a huge challenge in environments that are only modestly defined and structured. Furthermore, detection in the supply chain and "at the fork" should be also considered. This implies low-cost and power-autonomous compact sensors, connected to information processing systems used in the food supply chain and by consumers, that allow, for instance, freshness and food safety detection for meat and vegetables (which could be integrated into smartphones).

3.5.3.2.2 Crop Quality & Health

Integrated pest management (IMP)

Novel IPM strategies are needed to detect diseases and prevent their spread on crop production for European organic and conventional agricultures, to increase organic farming and allow the development

[26] "Our global food system is the primary driver of biodiversity loss," United Nations Environment Programme (UNEP), February 3, 2021

of horticultural systems that will use less/no pesticides. Improved IPM will require developments in the following fields:

- Smart systems based on portable real-time pest disease detection, diagnostics and monitoring platforms to provide rapid local and regional disease incidence alerts (georeferenced) e.g. weather/climate information for predictive models providing risk assessments and decision support for IPM.
- IoT devices specialized in pests and disease measurements, such as insect traps and other systems based on image recognition or AI models.
- Wearable Plant Sensors for directly monitoring the plant and being able to quickly detect pest appearance.

Agro ecology based: Move from conventional to organic, regenerative agriculture

To support the EU "*From Farm to Fork*" implementation, smart ECS can help farmers to drastically decrease the use of pesticides and their impact on human health and the environment. This will require:

- Development of cost-effective and intelligent intra-row, herbicide-free weeding techniques using advanced robots and robot fleets for individual plant recognition with high precision, based on advanced (vision) sensor technologies and edge AI working under in-field conditions.
- Development of smart and power efficient sensors to monitor the quality of spraying, as this is essential for biocontrol products and contact pesticides. Moreover, new sensors to monitor soil and plant health are needed such as pH, NO_3 & EC, soil moisture, CO_2, leaf wetness, surface temperature, airborne pathogens. They should be precise, low cost, highly miniaturized and biodegradable electronic and sensor components (printed antennas, organic batteries, biodegradable substrates, elimination of sensor/chip packaging) to avoid a negative Impact of electronic products on soil.
- Integration, into the same framework, of decision-support tools and precision agriculture tools to simplify farm management, improve crop quality and reduce costs.
- Advanced tools may yield on just farming 80% of the most productive soil leaving 20% for natural recovery reducing costs of fertilizers, energy to harvest and human time to collect, by this precision agriculture some complementary activities or species may be used to fix soil and reduce erosion and desertification.
- Dedicated services such as technical support (infrastructure deployment and maintenance), precision agriculture as a service, education, etc.

Plant precision breeding and plant phenotyping

The development of smart technologies can support precision plant breeding and phenotyping. This could be nanotechnology solutions or smart sensor solutions to support the following:

- Genomics and transcriptomics: DNA informed breeding, gene editing, genome prediction, breeding optimisation, phenotyping and seed sowing optimization.

- Large scale and high precision measurements of plant growth, architecture and composition: these measurements are required to optimize plant breeding by increasing our understanding of the genetic control and response of plants to their environment. Sensor systems should allow the study of plants in relation to biotic and abiotic factors, including plant-microbiome interactions, plant-plant competition, plant diseases and exposure to a multitude of variable abiotic environmental conditions such as light quality, irradiance levels, nutrient supply, temperature, humidity, soil pH and atmospheric CO2 levels. Plant health is also related to nutrients and the 4R strategy "right source at the right dose, at the right place, at the right time ", however, plant models, besides agri knowledge, are needed for nutrient content measurement and monitoring.

3.5.3.2.3 Livestock welfare and health

Livestock health is crucial for food safety. Different animal pathogens can be a serious threat for livestock management, because pathogens can be divided in those that can infect multiple species, and those that infect specific terrestrial and aquatic animal species (i.e., cattle, sheep, goats, equines, etc.). A great threat is represented by diseases that are highly transmissible, having the potential to spread rapidly across borders and cause significant socio-economic and public health consequences. All these lead to great economic loss to the farmers. Moreover, the use of antibiotics to treat animals increases the cost of production, as well as creating other problems like residues and resistance, which is of serious concern to public health.

In addition, better livestock management practices must reduce the release of pathogens into the environment and substantially reduce the risk of microbial contamination of surface and groundwater. As outlined in the literature[27], there is relevant impact of animal pathogens in four areas: animal health, economics, food safety and security, and public health. Consequently, there is strong interest in developing new advanced solutions systems with high sensitivity and specificity, for early detection of animal diseases and minimizing antibiotics use too.

Synergistic strategies can be applied to approach these challenges. Agronomic and technological solutions must be applied to:

- Minimise risks (e.g. release of pathogens into the environment, use of antibiotics for animals);
- Surveillance of the livestock environment through an efficient process of digitalization of livestock management to prevent diseases spread.

From a microbiological perspective, microorganisms can mitigate risk because specific beneficial microorganisms can be selected to work synergistically with other microorganisms already in the environment. Beneficial microorganisms can support the nutritional requirements of plants and reduce the incidence of pathogenic microorganisms, to solubilize minerals, to conserve energy, to maintain the

---

[27] T. A. McAllister, E. Topp, Animal Frontiers, Volume 2, Issue 2, April 2012, Pages 17–27, https://doi.org/10.2527/af.2012-0039

microbial-ecological balance of the soil, to increase photosynthetic efficiency, and to fix biological nitrogen.

Research has also studied beneficial effects when these specific microorganisms are included in animal diets[28]. It has been shown that when they meet the organic matter that makes up the animal's diet, they secrete beneficial substances such as vitamins, organic acids, chelated minerals, and antioxidants, influencing an antibacterial effect through a selective blocking of pathogen colonization.

Animal welfare is also an important concern for a growing number of consumers.

We can assess that these synergistic strategies (risk reduction with agronomic solutions and surveillance) as explained above are drivers for investing in better sensing systems for animal monitoring. Combined with data analytics solutions, this will improve animal health and welfare, resulting in more animal-friendly production, higher efficiency, better quality, and improved food control safety:

- Wearable sensors at the farm/barn level, and ambient sensors during cattle transport.
- Smart sensor systems to monitor animal activity, such as individual or group behaviour, to provide useful information for early detection of diseases and to increase animal well-being.
- Smart sensor systems for rapid verification of bacterial infection together with behavioural observations to control disease spread and support clinical and veterinary stakeholders to effect suitable therapeutic interventions; body temperature can also be monitored for early disease detection to reduce antibiotics use.

A major source of GHG emissions in agriculture is related to livestock enteric emissions that must be reduced to achieve the 1.5°C pathway. The methane emissions from livestock increase atmospheric temperature approximately 80 times more than CO_2 on a 20-year outlook, but methane has a shorter atmospheric lifetime than other GHGs, making it an effective target for reducing global temperatures quickly.

New methods are emerging to reduce enteric emissions in livestock, and these should be largely introduced, particularly in grassland or mixed systems, where cattle might be centrally handled only once or twice a year for weighing and treatment, and where their feed rations are unpredictable and uncontrollable. In addition to the smart sensors systems mentioned above, the following tools can be added:

- Methane verification and monitoring tools to measure methane emissions and sequestration.
- Besides GHG emissions, monitoring nitrogen emissions is crucial due to the significant impact these emissions have on the environment and public health. Nitrogen emissions primarily arise from the application of nitrogen-based fertilizers and the management of livestock, and they occur in various forms, including ammonia (NH_3), nitrous oxide (N_2O), and nitrates (NO_3^-). Smart

[28] D. Hidalgo, F. Corona & J.M. Martín-Marroquín Biomass Conversion and Biorefinery volume 12, pages 4649–4664 (2022)

sensor systems to monitor nitrogen emissions will play a pivotal role in addressing environmental challenges while optimizing agricultural productivity.

Taking into account the AI deployment in many sectors, the importance of a farmer-centric approach for livestock management is well highlighted within the digital livestock farming landscape[29].

When coupled with AI algorithms, all these data can provide real-time, objective, and holistic insights into animal welfare, but a farmer-centric approach is vital to ensure that technologies truly serve the needs of farmers and the welfare of animals. Therefore, every AI application should encourage this farmer centric approach, because the economic sustainability goal of the agricultural industry is also related to improve animal welfare and make informed decisions about disease prevention and management, leading to healthier herds and increased productivity.

3.5.3.2.4 Food Chain

Intelligent logistic systems including sensing and monitoring for food chains

Logistics are a critical component of the food chain. They not only determine the reach of distribution, but logistics delays and conditions profoundly affect the quality and safety of products reaching consumers and can result in food loss and waste in the supply chain.

Smart real-time sensing, monitoring and control systems in the food supply chain will safeguard food quality and food safety, while eventually reducing food losses in the supply chain. Therefore, technological solutions are required, but not limited to:

- Systems for monitoring and controlling the quality of food products and ingredients during transport and storage (e.g. temperature monitoring in cold chain, moisture, controlled atmosphere, ethanol, ethylene), which should be reliable, contaminant-free, secure, power efficient and interoperable along the logistics chain. Contact-less powering of the sensors placed in storage tanks could also avoid contamination of the food.
- Predictive systems to assess quality of (perishable) food products in the supply chain, providing real-time decision-support based on actual sensor measurements, supply chain data and AI models.
- Transport route optimisation, considering not only time and cost, but also external conditions and the intrinsic properties of the products being transported.

These needs are strongly related with traceability, as shown in the following section.

---

[29] Neethirajan, S. Artificial Intelligence and Sensor Innovations: Enhancing Livestock Welfare with a Human-Centric Approach. Hum-Cent Intell Syst 4, 77–92 (2024)

End-to-end food traceability

Food and beverage manufacturers and producers are faced with increasingly complex and fragmented supply chains, stricter regulation, and more demanding consumers. Regulatory compliance, competitive advantage, brand reputation and costs have made product traceability a priority and end-to-end traceability a major challenge. In today's globalised world where people of any origin live across every country, the source of food products and ingredients, as well as their certification, are a major concern and priority for consumers, who want to be sure about the origin of food products. Therefore, traceability should also encompass certifying food origin, as well as making information available on any relevant process to which the food product has been submitted. This information should not be restricted to mere tracking across the supply chain. End-to-end traceability solutions are required, but should not be limited to:

- Integrating blockchain into current technology to increase safety while preventing fraud and counterfeiting.
- Traceability to increase alignment between production and individual consumer demands, leading to better provisioning and more personalised nutrition support.
- Traceability to optimise distance between farm and fork – although many products are produced preferentially in specific parts of the world, there are also many examples of food that could be produced economically closer to consumers.
- Smart tags.

To this end, as IoT/IIoT solutions are increasingly being deployed, integrated hardware systems need to deliver (apart from mobility and connectivity) long lifetime autonomous sensing and AI-based intelligence at-the-edge, as well as edge and/or cloud analytics and cybersecurity, complying with privacy regulations where applicable, on a plug-and-play, open, interoperable architecture, and platform.

Distributed Ledger Technologies (DLTs) such as blockchain allow secure storage and tracking of all kinds of information, including condensed sensing or monitoring data regarding crops and livestock. Examples include information on crop seeds, feed ingested by livestock (including medication and antibiotics), as well as recording of the whole process that any farm product is submitted to until it reaches the consumer, throughout the respective supply chain and involved actors. Such information increases the transparency of these supply chains and can reduce potential production issues, e.g. simplify the tracing of eventual product spoiling, or other issues, to the respective source, supporting possible decisions to recall product batches if necessary. Consumers can also benefit from such transparency, i.e., they can be given access to information to make better informed decisions about the offered goods they want or need to acquire, and they can use the information to provide feedback to farmers and producers incentivising their policies and practices further.

Nowadays, end-consumers are more concerned about the origin of the agriculture products that they consume. A complete system to manage traceability and to offer, to the end-consumers, a complete transparency of the actions taken in the farm (for instance in the vineyard) and during the transformation process (in the example of a winery) is required: a solution could use blockchain technology. However, it is important that small-scale farmers with low technological expertise, resources, and insufficient size to integrate blockchain could be supported, eventually through a dedicated ecosystem. This can be devised through blockchains implemented to include such farmers and respective food product chains, which would also allow the support of food safety assurance, namely in respective local markets, eventually involving less costs associated with logistics and distribution, and thus also contributing to sustainability and fair food systems. Nevertheless, the lack of regulations, standardization and interoperability are challenges for incorporating blockchain.

3.5.3.3 Major Challenge 3: Environmental protection and sustainable production

3.5.3.3.1 Status, vision and expected outcome

EU regulations together with consumers' increased interest in organic food, is compelling farmers to drastically decrease the use of pesticides to reduce risks and impact on human health and the environment, as well as to undercut the maximum residue levels of pesticides. Pesticides are found not only in drinking water[30] but also in food and beverages. Lively debates have shown that our society demands alternatives to pesticides to help preserve the environment and improve food quality.

Drastic reduction in the use of pesticides is one of the major goals of the EU's agricultural policy, with some countries planning to halve their pesticide use by 2025 (e.g. ecophyto plans[31] in France, and the Aktionsplan Pflanzenschutzmittel[32] in Switzerland). The EU Farm-to-Fork strategy also aims to implement an action plan that significantly reduces risks from chemical pesticides, as well as the use of fertilisers and antibiotics, and to increase the amount of organic farming carried out in Europe.

In general, the new EU agriculture policies put a major focus on preservation of landscapes, biodiversity, and environmental protection, in a results-oriented model aligned with the Green Deal[33]. For instance, the reform of the Common Agriculture Policy (CAP) introduces measures for fostering the adoption of sustainable farming practices ("eco-schemes"), such as agroecology or organic farming. Farmers will need to provide "digital evidence" of compliance to the CAP rules and the implementation of good practices. The CAP evaluations will be largely based on the use of high-quality data collected directly from the field. This will require measurement and monitoring technology (for environmental performance, biodiversity monitoring parameters, etc.) which is accurate, highly scalable, and secure (certified monitoring information).

Areas of interest are often remote without sufficient connectivity – new approaches are required to flexibly deploy sensors that harvest their energy and collect data.

---

[30] https://www.notre-environnement.gouv.fr/themes/sante/la-pollution-de-l-eau-douce-ressources/article/les-pesticides-dans-les-eaux-souterraines?lien-ressource=5193&ancreretour=lireplus

[31] Ministère de l'Agriculture, Le Plan Ecophyto, qu'est-ce que c'est? https://agriculture.gouv.fr/le-plan-ecophyto-quest-ce-que-cest , 2020

[32] Aktionsplan Pflanzenschutzmittel, https://www.blw.admin.ch/blw/de/home/nachhaltige-produktion/pflanzenschutz/aktionsplan.html

[33] https://aioti.eu/wp-content/uploads/2022/02/AIOTI-Role-of-IoT-in-addressing-agroecological-focus-of-Green-Deal-Final.pdf

3.5.3.3.2 Soil Health

***In-situ* real-time monitoring of soil nutrients and herbicides**

The optimal use of chemical fertilisers and organic manures to deliver the ever-increasing food production requires a complete understanding of the nitrogen- and phosphorous-based nutrients applied in the fields with a much greater spatial and temporal resolution than is available today. Current methods of soil analysis do not provide real-time, precise and *in situ* nutrient analysis in fine detail, and delays in receiving soil results are common because of backlogs in commercial labs due to high sample volumes, thus reducing the value of the soil test results for the farmer. Moreover, herbicide application is another huge problem due to their environmental and health impact. To solve these issues, the following actions must be done:

- Intelligent sensors and bio-sensors (with miniaturised and ultra-low power consumption components allowing these sensors to harvest their own energy) must be developed to deliver measurements of soil quality and soil nutrients in situ and in real time at parts per million (ppm) concentrations. Such devices must have the appropriate packaging to extract water from the soil. Ideally, they should be buried in the soil for long periods of time or at least while sustaining operation capabilities for the entire growing season. To optimise effectiveness, low proximity sensors should be combined with optical sensors and high proximity sensors to retrieve the maximum amount of information on soil health.
- Likewise, smart actuators could prove to be highly beneficial. Such miniature units could be deposited on or buried into the soil. Coupled to sensing functionalities into the same module, critical actions (e.g. release of agrochemicals) could be triggered very close to the plant roots for maximized efficiency.
- Multidisciplinary approaches for developing novel sustainable smart ECS are needed. Indeed, current ECS contain a variety of toxic materials and chemicals. As such, they cannot be left in the soil. The optimization in the use of agrochemicals should not come at the expense of another ecological burden. New sustainable "green" ECS made of eco-friendly materials that will have benign environmental impact must be created.

IoT systems with edge and/or cloud-based data analytics are also necessary to provide farmers with decision-support regarding fertilisation strategies, by translating measurements into meaningful agronomic indicators and respective measures. These strategies should prioritise the use of organic fertilisers and the gradual reduction of chemical ones until eliminated to restore the biodiversity contribution in the preservation of soil health. Furthermore, this type of system should detect weeds, preserve the “good ones” and eradicate those that are competing with the crop in question. This requires low-cost vision technologies (not only plain optical red/green/blue (RGB), but also 3D, hyperspectral imaging, etc.) and edge AI for in situ prompt recognition and decision-making.

3.5.3.3.3 Healthy air and skies

Sensors and diagnostics for air quality monitoring (indoor, urban and rural)

According to the World Health Organization, the air we breathe is becoming dangerously polluted. Nine out of ten people now breathe polluted air, which kills seven million people every year. There has been much progress on identifying and reducing the sources of air pollution at lower concentrations and with higher spatial coverage. This is necessary to provide adequate data on what people are breathing, and to provide localised as well as holistic solutions. Microsensors and/or mini-stations can be used during fieldwork campaigns in cities, but there are technical problems relating to power source, data transmission, data storage, and data handling and assessment. Besides, local measures are not always effective since local concentrations of particulate matter may be influenced by long-range transported pollutants from agricultural activities occurring outside city boundaries.

Similarly, while indoor air quality has been shown to unambiguously impact the wellness, health and performance of people as shown for instance with Covid-19 in schools due to lack of indoor air quality measurement, there is also a lack of spatial granularity and a significant lag between exposures and sensing, actuation and management interventions for risk mitigation. In addition to indoors, air quality is made more complex by the interaction between indoor and outdoor air, emissions from buildings and their contents (paints, furniture, heating, and cooling systems, etc.), human activities (breathing, cooking, cleaning, etc.) and the effects of long-term exposure to low concentrations of volatile organic compounds. These issues necessitate development and deployment of real-time intelligent multi-sensor technologies with high selectivity and embedded (re-)calibration techniques. These should be combined with a monitoring network (edge-based) as part of the indoor infrastructure to provide the spatial and temporal information needed for specific, targeted and appropriate actions. A required ingredient of this monitoring solution is a digital twin platform able to predict in real-time the pollutant diffusion, considering the complex environment monitored. Such actions should also include public awareness and the promotion of behavioural changes.

Smart systems for controlling and preventing GHG emissions

Strong evidence has been accumulated on the climate emergency resulting from human activities that add GHGs to the Earth's atmosphere. The EU is the world's third biggest GHG emitter after China and the US. Although several measures have been taken since the Paris Agreement, breakthrough technologies and state-of-the-art deployment are still needed across the transport sector and other industries with a high emission footprint to achieve a further reduction in emissions. These would be facilitated by the following:

- Smart systems and digitalisation to improve industrial processes performance and energy/resource efficiency towards a low-carbon economy, while reducing the impact of mobility and agricultural processes on the environment and human health, thereby controlling and preventing GHG emissions.
- A focus on the GHG emissions from animals by investigating microbiological sensing technologies on or in animals (in their rumens or breath, for instance) to increase efficiency while reducing environmental impact, as well as performing analysis of the gathered data to support decision-making for mitigation measures (for instance, leading to change in feed).

- GHG verification and monitoring tools to measure GHG emissions and sequestration.

3.5.3.3.4 Smart Waste Management

Integrated waste systems

Despite proactive European policies and regulations[34], effective bio-waste management remains a challenge. Reducing, recycling and reusing food/kitchen waste requires significant progress in technological solutions along with strong policymaking and shifting community behaviour. These solutions could be based on the following:

- Smart monitoring, controlling waste treatment units in real-time as well as gas emissions in landfills and anaerobic digestion monitoring. Data analytics should include gamification for behavioural triggers.
- Smart real-time quality control systems of the sorted waste (i.e., measure of the "purity" of the sorted waste) to ensure their proper recyclability
- Smart waste collection bins (radio-frequency identification (RFID) tags, self-compacting bins, fullness level sensors, automated waste segregation), even in remote locations, and without access to power supply, including automated robotic systems and optimised separation systems, which can be complemented by the upcycling of waste streams into usable resources and optimal routing systems, as well as vehicle tracking. These solutions should be integrated and interconnected into the product life cycle "from cradle to grave" to enable circular and resource-efficient methodologies.

Intelligent sustainable / biodegradable packaging

Intelligent and biodegradable packaging concepts have been gaining traction in the food industry to improve product safety and reduce environmental impact. Smart sensors in an IoT system can monitor environmental conditions and product quality, while communication devices can store and convey data throughout the product life cycle. While these concepts need to be further advanced for efficient, safe food production and waste management, intelligent packaging itself needs to become more sustainable. Novel ideas are required to solve the problem of the amount of plastic packaging produced by food manufacturers. The definition of biodegradable packaging should lead to a new generation of food packaging. Such novel ideas include:

- A synergetic interdisciplinary approach to cross the boundaries of novel materials for food packaging and smart sensors associated with analytical methods for the detection of harmful substances that can infiltrate into food, cause water contamination, etc.
- Fabrication and hybrid integration of eco-friendly nanostructured electrodes, sensors, energy harvesting and storage devices on rigid and flexible biodegradable substrates to reduce the waste from embedded electronics in smart packaging.

---

[34] https://www.consilium.europa.eu/en/press/press-releases/2018/05/22/waste-management-and-recycling-council-adopts-new-rules/

3.5.3.3.5 Remediation

Efficient smart networks for remediation

Remediation processes aimed at converting harmful molecules into benign ones can be undertaken in different ecosystems, such as water bodies (e.g. biotic, and abiotic farming by-products), air (e.g. GHGs) and soil (e.g. pesticides). Remediation processes are mainly carried out in wastewater treatment plants. Although some pollution sources are static and sufficiently well-known such that treatment can be undertaken effectively, other pollution sources are more mobile in both time and/or space, making treatment at single points unsatisfactory. Another limiting issue is that remediation technologies are often power-intensive and cannot be deployed for long in remote locations. Alternative high-efficiency remediation methods are needed, such as to transform/reduce the levels of CO2 in chemical products. Current devices are also prone to fouling. This means remediation processes cannot be run constantly in remote locations, and there is thus a necessity to undertake them only when and where they are most required. In this regard:

- A network of smart sensors (an IoT system) that can monitor relevant status in real time, and inform on the necessity of remediation, would provide unique decision support invaluable for efficient water, air, and soil management.
- Techniques used in the measurement and analysis of carbon sequestration by soils could also investigate the current potential of soils as a remediation mechanism to improve the sequestration capacity – such investigation should include the initiative of “four per 1000”[35] presented at COP21 in Paris.
- Likewise, tools and methods able to evaluate the performance of the carbon sequestration techniques employed should be developed to guarantee their efficiency.

3.5.3.4 Major Challenge 4: Water resource management

3.5.3.4.1 Status, vision and expected outcome

The quality of groundwater, surface water bodies (oceans, seas, lakes), waterways (rivers, canals, estuaries) and coastal areas has a great impact on both biodiversity and the quality of water that people consume every day. While natural droughts may lead to increased salinity in freshwater systems[36] and along with floods, impact or endanger the quality of water bodies, human activities in energy production, data centers, manufacturing and farming industries have a major detrimental effect through thermal

---

[35] Researchers from the French National Institute for Agronomic Research, mentors of the project, have observed that by increasing the organic matter in the soil by 4 grams per 1000 - hence its name - it would be possible to limit the current growth of CO_2 emissions to the atmosphere. Promoting good agricultural practice would combat climate change and, at the same time, guarantee the food security of the population.

[36] https://www.sciencedirect.com/science/article/abs/pii/S0012825214002086

pollution, chemical, microbiological and micro-plastic contaminants, and biotic and abiotic farming by-products. These human activities have released pollutants of emerging concern (such as PFAs and medicinal products) into water bodies, with detrimental consequences for human health and biodiversity. Moreover, the outdated and deteriorating water infrastructure is also having a detrimental impact on both water quality and the amount of water lost through leakage.

In the context of climate change, increased water temperatures may cause (apart from extreme evaporation) eutrophication and excess algal growth in surface water bodies. Moreover, heavy storms increase the amount of sediment nutrients and pollutants in water sources, and human activities release pollutant of emerging concerns that end up in all types of water bodies which have a direct impact on drinking water quality. Therefore, climate change jeopardizes the quality and safety of our water, making the development of new tools to deal with this problem more critical than ever.

3.5.3.4.2 Access to clean water (urban and rural)

Healthy Water

With the aim of reducing pollution-related health problems, water utilities, water associations, academia and private industry have focused on developing new methods, policies and procedures to secure drinking water distribution by (1) detecting in real-time any compound, contaminant or anomaly that may represent a health risk for the end-users and (2) taking the required measures to mitigate these issues. This necessitates online information on the status of water sources at a scale larger than ever before. To mitigate both accidental and intentional contamination of freshwater sources, the deployment of sensors and diagnostic and decision support systems with rapid communication technologies and data analysis capabilities are needed to secure water quality and its distribution over the network. Such actions shall provide:

Connected and highly integrated low power (or self-powered for maintenance/battery-free system) multi-parameter diagnostic sensors for real-time physico-chemical analysis (temperature, ionic electrical conductivity, pH, turbidity, inorganic pollutants as nitrates or heavy metals, etc.) in water distribution network and wastewater treatment plants, and biofilm growth monitoring in water pipes. Online monitoring systems at the edge, including devices with embedded AI for data analytics. The presence of a mesh of intelligent devices in drinking water networks will make it possible to identify and deal very quickly with drifts or anomalies (e.g, leaks, contamination events, etc.) while reducing the amount of data sent to the servers. Prediction in real-time of pollutant diffusion with simplified compact models, considering non-dense measurements from the multi-parameter diagnostic sensors and online monitoring that are developed and deployed.

Integrated systems for demand reduction and conservation of water

According to the UN Development Programme[37], dwindling drinking water supplies is affecting every continent. On the one hand, increased urbanisation and farming have amplified the demand of water for

[37] https://www.undp.org/

human consumption and for domestic and agricultural use. On the other hand, an increasing number of countries are experiencing water stress due to longer drought periods and the spread of desertification. In addition, approximately 25% of all urban drinking water is being lost forever[38] in global water distribution systems, before it even reaches the end-user. Therefore, there is an urgent need to prevent losses from water abstraction as climate effects intensify. Leak localisation is currently very time-consuming, labour-intensive, and costly. Operators must manually place equipment that "listens" to the water flow during the night. Smart integrated systems can significantly contribute to key measures aiming at affecting consumer practices in water usage, delivering greater efficiency in detecting leaks and ultimately reducing water waste. Developments are needed in the fields of:

- Smart metering, time-of-use pricing and gamification to change habits in water consumption and control appliances, along with interoperable solutions for a truly connected smart household (taps, lavatories, showers, appliances).
- Low-cost self-powered/ low-power sensor nodes for flow control, leak detection and auto shut-off, along with inexpensive actuators to remotely control valves for limiting water usage by volume/time. IoT systems can optimise the control of household, agricultural and industrial infrastructure/equipment in water-intensive processes.
- Smart systems able to automate leak localisation, and to respond promptly and cost-effectively. This can be a combination of in-pipe inspection (to locate the leak) and a network of low-cost, fine-grained sensors to allow predictive maintenance of distribution systems.

Efficient and intelligent water distribution

The main challenge for improving the use of water is to guide its distribution depending on its final application (drinking water, water for industry, water for the cooling of data centers, etc.). However, the existing sanitary regulations always look to optimise water safety regardless of its final use. To apply the most effective measures to make water distribution more efficient, it is necessary to thoroughly review the different supply protocols and quality criteria for each sector. Moreover, by continuously monitoring the quality and availability of water, it would be possible to better regulate its distribution depending on the final use and to adjust the price accordingly. Intelligent systems connected to smart grids will allow water inputs to the network to be made at the right times, optimising the energy cost as a result.

To address these challenges, there is a need for developing:

- Novel smart metering solutions that include electrochemical multi-parameter sensors (pH, chlorine, conductivity, etc.) with high stability, anti-fouling, high accuracy capabilities and cost-effectiveness, as well as optical sensors based on different principles (fluorescence, absorbance, etc.) integrated into miniaturised systems at a low cost.

---

[38] Leakage Reduction in European Water Mains; Layman Report https://ec.europa.eu/environment/eco-innovation/projects/sites/eco-innovation-projects/files/projects/ documents/-curapipe_layman_report.pdf

- Robust IoT systems with adequate data analysis processing power and AI capabilities to handle the large volume of data generated by the different water management processes to satisfy quality, usage type and associated pricing.
- Efficient year-round water management in terms of storage to deal with some of the most urgent shortages, with better forecasting and warning systems based on extensive measurements – e.g. intentionally flooded areas could be used to store water in times of expected scarcity.
- Prediction of pollutant diffusion in water distribution systems.

3.5.3.4.3 Resource Management

Smart systems for irrigation management

At a global level, agriculture consumes 69% of the world's freshwater[39]. Because of this, precise control of irrigation is essential to guarantee water and food security for all. Irrigation water management is the practice of monitoring and managing the rate, volume, and timing of water applications according to seasonal crop needs, considering the soil intake and water holding capacities with the objective of using water in the most profitable way at sustainable production levels. To this end:

- Smart sensors, with low-cost and self-powered characteristics, are increasingly required as tools to implement irrigation management and monitor water levels. Sensors should be more intelligent to support real-time applications and/ or reduce latency. Optimisation of the power consumption of the overall system shall be done to enable ultra-low power, ideally self-powered battery-free solutions that are cost effective and seamlessly deployed at the edge for both outdoor and indoor use.
- Integration of systems monitoring water deficiency or surplus is also required. These could be based on narrow-band spectral reflectance of water and land surfaces for vegetation/habitat mapping, along with UAV utilisation in remote areas.
- Smart low cost actuators will provide the "right dose (of water) at the right place".
- Appropriate simplified models should allow to limit the number of sensors spread over a given landscape and integration of various sources of data, included satellite ones, will expand the analysis capability of decision support systems. Also, leak detection in the irrigation management system shall be considered to ensure that water is not lost.

Smart systems for flood management

Flood management has been gradually integrating smart sensors. IoT systems with water-level sensors can also play a significant role in real-time monitoring and natural hazards predictive/forecasting capacity models. This requires:

---

[39] http://www.fao.org/nr/water/aquastat/water_use/

- The monitoring of water levels and devising prediction models to identify areas at a high risk of flooding. This is possible through the development and deployment of more intelligent low power sensor nodes in combination with smart predictive algorithms that will act as digital twins to integrate information from various sources, such as weather forecasts and regional georeferenced data.
- IoT interoperable systems are key for provision of real-time information to first responders, civilians and companies to proactively take countermeasures.

Smart water treatments fostering circular use (wastewater, rainwater, storm water)

Around 80% of all wastewater is currently being discharged into the world's waterways, where it creates health, environmental and climate-related problems. Water from industrial, agriculture and domestic use contains organics, phosphates, nitrogen, cellulose, rare earth elements and other substances (including residues of medicinal products. In addition to its domestic use, purifying, distilling, or deionising water is essential for many agricultural and industrial uses – to ensure the consistency of products and to meet strict safety regulations. The global market for water and wastewater technologies reached US $64.4 billion in 2018 and is expected to rise to US $83 billion by 2023[40]. Commercial technologies that allow resource recovery from wastewater to be commercially feasible are increasingly being developed, making transitioning to a circular economy an opportunity to accelerate and scale-up the most recent scientific and technological advances that support greater efficiency in the water sector. However, this requires further advances in technologies such as:

- A range of sensors in water systems to monitor water levels, the flow of water through different channels, temperature changes, chemical leakage, pressure level, chemical residues, and biological residues, pH, etc., associated to smart decision support systems that allow to monito, take action in real-time and forecast the water treatment plant, considering multiple sources of data (from the sensors deployed on-site but also e.g., weather forecast).
- IoT-enabled water purifiers that can predict potential system failures to reduce downtime in water treatment plants, and to enable remote sensing for mapping groundwater resources and monitoring sustainable extraction levels.
- IoT-enabled increase of water recycling and development of the urban circular water economy. As examples waste water from industry and data centers can be used for district heating or treated and recycled to yield not only water but also energy, fertilizer, and organic inputs.
- IoT-enabled increase of water recycling in water-demanding industries (e.g., mining, semiconductor manufacturing) with reduction of the water footprint in manufacturing. Water-efficient processes must be implemented such as closed-loop systems, which capture and recycle the water used in the production process.

---

[40] https://www.bccresearch.com/market-research/environment/water-and-wastewater-treatment-technologies-global-markets.html

3.5.3.5 *Major Challenge 5: Biodiversity restoration for ecosystems resilience, conservation, and preservation*

3.5.3.5.1 Status, vision and expected outcome

It has been stated that: "Biodiversity boosts ecosystem productivity where each species no matter how small, all have an important role to play"[41]. For example, increasing the number of plant species means a greater biodiversity ensuring natural sustainability for all life [42]. Healthy ecosystems can better withstand and recover from a variety of disasters, anthropogenic or not. Healthy biodiversity offers many natural services for everyone.

It should be noted that there are many such services that we already get for free! However, the cost of replacing these, even, if possible, would be extremely expensive. More than ever, as noted in Section 3.5.4.3, the new EU agriculture policies promote sustainable farming practices that help to protect the environment, preserving landscapes and biodiversity. This is a consequence of the well-recognized correlation between the health of ecosystems and the health of farming production. It therefore makes economic and development sense to move towards sustainability. From this perspective, ECS will contribute to addressing some of the key challenges relating to biodiversity and sustainability for the four ecosystems described below.

3.5.3.5.2 Biodiversity restoration for the agriculture ecosystem

Among the key focus areas Agriculture is one of the economic activities that has the highest dependence on nature and biodiversity[43]. On average, global mean crop yields of rice, maize and wheat are projected to decrease between 3% and 10% per Celsius degree of warming above historical levels. All crops depend directly on soil health and fertility, and more than 75% of global food crop types rely on animal pollination. However, the impact of agriculture activity on the environment must be as low as possible to preserve biodiversity. Efforts to conserve existing land resources (e.g. forests) and expand natural-based solutions (e.g. peatlands restoration) are required to reduce the GHG emissions, or to improve resistance by microbial biofertilisers [44] [45].In this regard, the EU Biodiversity Strategy 2030 establishes several objectives[46], summarized in *sub-section 3.5.5* Timeline. To address these objectives, there is a need to develop:

---

[41] Anup Shah « Why is biodiversity important? Who cares?
https://www.globalissues.org/article/170/why-is-biodiversity-important-who-cares

[42] https://www.pnas.org/doi/10.1073/pnas.2203385119

[43] European Commission. The business case for biodiversity. May 2020.
https://ec.europa.eu/commission/presscorner/detail/en/fs_20_907

[44] https://www.fao.org/3/nd651en/nd651en.pdf

[45] Prisa, D, Fresco, R, Spagnuolo, D. 2023. Microbial Biofertilisers in plant production and resistance: a review. Agriculture 13(9), 1666 - MDPI

[46] European Commission. EU Biodiversity Strategy for 2030: bringing nature back into our lives. May 2020.
https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1590574123338&uri=CELEX%3A52020DC0380

- Precision farming systems and services for optimal use of fertilisers and pesticides.
- Sensing and monitoring systems for *in situ* measurement of soil nutrients, connected insect traps and landscape monitoring.

3.5.3.5.3 Biodiversity restoration for the aquaculture ecosystem

Aquaculture impacts biodiversity negatively in several ways[47]: (i) where antibiotics and hormones are used to reduce farm stock mortality and improve growth rates, but their use has side effects for the flora and fauna of water bodies receiving farm effluents; (ii) through eutrophication and changes in flora and fauna in waters receiving effluents from aquaculture facilities; (iii) through the risk of excessive exploitation of wild fish stocks for use in farm fish feeds; and (iv) by transfer of disease and parasites from farm animals to wild animals.

To address these side effects, there is a need to develop:

- Precision aquaculture systems for optimal feeding (minimizing waste and feed residuals), optimal use of antibiotics/hormones, and optimal use of freshwater.
- Smart multi-sensors and smart systems for monitoring water quality in aquaculture facilities and their effluents.
- Smart systems combining data collected from different sources (IoT networks, satellite, and drones) and data analysis based on AI/ML techniques to create predictive models leading to more confident decision-making, timely alerts, and automated systems in general.

3.5.3.5.4 Biodiversity restoration for the fisheries ecosystem

The EU's Biodiversity Strategy has set an objective of protecting a minimum of 30% of its sea area. Like agriculture, fishing is an economic activity with a strong dependence on biodiversity. Keeping fish stocks healthy is critical to guaranteeing ocean biodiversity and thus the economic sustainability of fisheries. According to this strategy, the preservation of marine stocks could increase the annual profits of the European seafood industry by more than €49 billion.

Fishing activities impact biodiversity negatively in several ways, particularly by: (i) increasing fish mortality, so measures must be taken to keep this under maximum sustainable yield levels; and (ii) damaging the ocean ecosystem due to the use of certain fishing techniques, currently the most damaging activity to the seabed. In addition, the effect of by-catching from non-selective industrial fishing methods endangers many species of marine animals not being fished for. It is therefore necessary to evolve towards more selective and less damaging fishing techniques, as well as the more effective control of illegal fishing practices.

---

[47] Claude E. Boyd, What is biodiversity and its relevance to aquaculture certification? https://www.aquaculturealliance.org/advocate/biodiversity-relevance-aquaculture-certification/

To reduce these negative impacts, there is a need to develop:

- Oceanographic sensing and monitoring solutions (including unmanned vehicles, UXVs) for fisheries ecosystems to estimate biodiversity indices, fish stocks and species distribution, and to build fishery management systems consistent with conservation objectives and rules, also related to the tropicalization process that it is becoming emergent in last times.
- Technologies to make fishing gear more selective and environmentally respectful.
- Technologies for checking compliance and detecting illegal activities (onboard cameras, RFID, traceability technologies, vessel monitoring, etc.).

3.5.3.5.5 Biodiversity restoration for the forestry ecosystem

The EU Biodiversity Strategy has set the objective of protecting a minimum of 30% of the EU's land area. At least one-third of protected areas – representing 10% of EU land – should be strictly protected. In particular, the strategy identifies the crucial need to strictly protect all the EU's primary and old-growth forests, which are the richest forest ecosystems removing carbon from the atmosphere, while storing significant carbon stocks.**Error! Unknown switch argument.** The strategy also calls for preserving the good health and increasing the resilience of all EU forests, especially against wildfires, droughts, pests and diseases. It is envisaged that the European Commission will develop a forest information system for Europe that integrates data from multiple sources and providers. To prevent more wildfires, we need to grow rural economies in a sustainable way and manage climate change, with a much better understanding and continuous assessment of EU forests. To this end, there is a need to develop:

- A precision forestry system with remote self powered sensing and AI/ML monitoring capabilities to map and assess the condition of the EU forests as well as early detection and prevention of threats to the forests (wildfires, pests, diseases, etc.).
- Smart systems for environment monitoring of forests as well as C02 footprint monitoring, remote monitoring of wildlife behaviour and habitat changes, and provide timely warning on illegal poaching activity.
- Customised services (similar to precision agriculture as a service as discussed earlier), not only to support the above-mentioned systems but also to further exploit the information they provide.

3.5.4 TIMELINE

MAJOR CHALLENGE	TOPIC	SHORT -TERM 2025 - 2029
Major Challenge 1: food security	**Topic 1.1:** intelligent and adaptative food production	• Advanced analytical processing based on several data sources. • IoT devices with integrated firmware for implementing big data solutions

	Topic 1.2: redesigning farming systems	• Advanced autonomous robotic systems and small robots for labour free, ecological friendly farming including smart sensors, edge AI • A farm management information system (FMIS) with decision support thoroughly integrated with IoT and automated systems; all the digital data should be gathered automatically
Major Challenge 2: food safety	**Topic 2.1:** crop quality and health	• IoT for monitoring the key parameters related to plant health including hydric stress. • Decision Support Systems (DSS) for recommendation/decisions related to agrochemical application; health and environmental care
	Topic 2.2: livestock welfare and health	• Advanced indicators of welfare, health and performance monitoring (integration of milking robot, wearable sensors, etc.) at the individual and herd scale • IoT devices (sensors) for monitoring livestock emissions of GHG, nitrogen
	Topic 2.3: food chain	• IoT devices monitoring food quality, safety and transport from production to the retailer; end-consumers to have full access to this information; AI (ML/deep learning) models based on the recommendations and decisions that the IoT devices could take to monitor the whole supply chain • Global accessibility for end-consumers to the traceability of the whole value chain – i.e. total transparency
Major Challenge 3: environmental protection and sustainable production	**Topic 3.1**: soil health	• Autonomous recommendation system related to fertilisation and phytosanitary application, considering measurements from multi-parameters IoT devices and other sources of information (e.g. weather forecast). • Sensor system to measure fertiliser content in manure prior to their application in fields
	Topic 3.2: healthy air and skies	• CO_2 capture materials in use • Advanced sensors for air quality (e.g. particular matters, GHG, nitrogen)
	Topic 3.3: smart waste management	• Forecasting models of potential waste that will be produced by the farm management system
	Topic 3.4: remediation	• Network of sensors for target pollutant with antifouling properties for use in real environments • Development of capture materials for targeted pollutants, including CO_2 capture materials

Major Challenge 4: water resource management	**Topic 4.1**: access to clean water (urban and rural)	• ICT solutions allowing greater societal involvement in water management through online knowledge of its consumption data (remote meter reading), and quality parameter monitoring for greater awareness about the optimisation of the freshwater as a limited resource • Water quality monitoring systems based on hybrid technology (mono-parameter bulky probes and some miniature chips) • Multi-parameter sensor solutions for *in situ* real-time measurement (included in remote locations): Sensors for basic parameters such as chlorine, conductivity and pH are available for real-time monitoring; more complex parameters require lab analysis • Cost and integration are still challenging for massive deployment in water distribution networks based on current IoT system applications • Limited amount of data (systems are installed only at critical locations) • Centralised control and data analysis based on AI on the cloud
	Topic 4.2: resource management	• Requirements identification and classification for biodiversity protection in the exploitation of aquifers for human supply • Monitoring systems for the water lifecycle, including supply and sanitation through the development of multi-parameter sensor nodes and digital tools allowing the intensification circular economy • Progressive transformation of wastewater into raw materials for the generation of products and services
Major Challenge 5: biodiversity restoration for ecosystems resilience, conservation and preservation	**Topic 5.1**: biodiversity restoration for agriculture ecosystem	• Sensing and monitoring systems for in situ real-time measurement of soil nutrients , connected Insect traps and landscape monitoring
	Topic 5.2: biodiversity restoration for aquaculture ecosystem	• Smart multi-sensors and smart systems for monitoring water quality in aquaculture facilities and their effluents
	Topic 5.3: biodiversity restoration for fisheries ecosystem	• Technologies for checking compliance and detecting illegal activities (onboard cameras, RFID, traceability technologies, vessel monitoring, etc.)
	Topic 5.4: biodiversity restoration for forestry ecosystem	• Precision forestry system with remote sensing and AI/ML monitoring capabilities to map and assess the condition of the EU forests, as well as early detection and prevention of threats to forests (wildfires, pests, diseases, etc.)

MAJOR CHALLENGE	TOPIC	MEDIUM-TERM 2030-2034
Major Challenge 1: food security	**Topic 1.1:** intelligent and adaptative food production	• AI applied to food production to define advanced analytical processing related to prescriptive and predictive analysis
	Topic 1.2: redesigning farming systems	• Semi-autonomous agronomic systems (irrigation systems, climate control systems, etc.) based on expertise and farmers' decision-support systems (DSS)
Major Challenge 2: food safety	**Topic 2.1:** crop quality and health	• AI for decisions and action support with self adaptation and learning capabilities; ML and deep learning related to agronomic models and algorithms
	Topic 2.2: livestock welfare and health	• Technoogy to enable the reduction in the use of antimicrobials for farmed animals by 50% by 2030
	Topic 2.3: food chain	• Interoperability among all the systems that manage the whole value chain • Normalisation and homogenisation of communication protocols and end-to-end security • IoT devices integrated in the food chain where the end-consumers will be able to read them by mobile phone and directly access for complete traceability
Major Challenge 3: environmental protection and sustainable production	**Topic 3.1**: soil health	• Combination of several data sources to establish and attain key performance indicators (KPIs) related to environmental protection and sustainable production
	Topic 3.2: healthy air and skies	• CO2 capture and conversion on site
	Topic 3.3: smart waste management	• Registration of the traceability related to residues management, including the residue management in food traceability and the environmental footprint
	Topic 3.4: remediation	• Coupled sensor and CO_2 capture/conversion system for CO_2 remediation • Solar/thermoelectric in situ driven pollutant removal
Major Challenge 4: water resource management	**Topic 4.1**: access to clean water (urban and rural)	• Smart monitoring systems at home to optimise household water spending and tools to improve performances through KPIs that allow for measuring progress at the microscale; water users must move from passive consumers to active management • New generation of more integrated and miniaturised multiparameter autonomous sensors (e.g. pH, chlorine, and conductivity parameters)

		• More complex sensors are available for real-time detection of pollutants in water, such as heavy metals and nitrates • Edge computing and multiparameter devices allowing decentralised data analysis and control • Massive deployment starts being cost-effective with more accurate solutions due to the availability of an increased amount of data
	Topic 4.2: resource management	• Improvement of knowledge through the accumulation of consolidated and valid data series, on the natural environment through the implementation of monitoring systems, for both the water and natural environment (fauna, ecology, sociological aspects, uses, etc.), as a basis for sustainable management through AI/ML tools, allowing for identification of the correlation between the evolution of the environment quality and water use • Design of environmental evolution models in different use scenarios • Industrial transformation of wastewater treatment plants in bio-factories
Major Challenge 5: biodiversity restoration for ecosystems resilience, conservation and preservation	**Topic 5.1**: biodiversity restoration for agriculture ecosystem	• Precision farming systems for optimal use of fertilisers and pesticides • Reduction of the use and risk of chemical and more hazardous pesticides by 50% by 2030 • Reduction of nutrient losses by at least 50% while ensuring no deterioration to soil fertility • Reduction in fertiliser use by at least 20% by 2030 • Reduction in the sales of antimicrobials for farmed animals and in aquaculture by 50% by 2030 • Boosting the development of EU organic farming areas to achieve a 25% increase in total farmland under organic farming by 2030
	Topic 5.2: biodiversity restoration for aquaculture ecosystem	• Smart systems combining data collected from different sources (IoT, satellite and drones) and data analysis based on AI/ML techniques and digital twin to create predictive models leading to more confident decision-making, timely alerts and automated systems in general
	Topic 5.3: biodiversity restoration for fisheries ecosystem	• Oceanographic sensing and monitoring solutions (including UXVs) for fisheries ecosystem to estimate biodiversity indices, fish stocks and species distribution
	Topic 5.4: biodiversity restoration for forestry ecosystem	• Smart systems for environmental monitoring of forests and fields, as well as CO_2 footprint monitoring, remote monitoring of wildlife behaviour and habitat changes, and provision of timely warnings about illegal poaching activity

MAJOR CHALLENGE	TOPIC	LONG-TERM 2035 AND BEYOND
Major Challenge 1: food security	**Topic 1.1:** intelligent and adaptative food production	• AI applied to food production, not only in pre-harvest areas but also post-harvest – i.e. applied to the whole value chain integrally
	Topic 1.2: redesigning farming systems	• Automation of labour; resource optimisation (further targeting environmental care and social impact)
Major Challenge 2: food safety	**Topic 2.1:** crop quality and health	• Robots with AI for managing plant health autono-mously
	Topic 2.2: livestock welfare and health	• Fully automated herd performance control (growth and milk production, forage efficiency, early disease detec-tion for antibiotics use reduction), and applications for genetic selection to optimise breeding performance and resilience
	Topic 2.3: food chain	• IoT devices making recommendations automatically and take autonomous decisions related to food safety, acting directly with the transport mechanism (cooling mechanism and others that impact food safety) • Systems automatically and autonomously act in all the machinery located at each step of the supply chain
Major Challenge 3: environmental protection and sustainable production	**Topic 3.1**: soil health	• Autonomous actions performed by IoT devices directly in systems related to fertilisation and phytosanitary ap-plications
	Topic 3.2: healthy air and skies	• Low or no carbon fuel sources
	Topic 3.3: smart waste management	• AI and digital twin models providing recommendations for decision-making related to minimising farms waste
	Topic 3.4: remediation	• Real-time multiparameter sensing with AI and digital twin decision-support for management • Efficient and low-cost general pollutant removal and conversion systems using energy harvesting towards in situ remediation
Major Challenge 4: water resource management	**Topic 4.1**: access to clean water (urban and rural)	• Use of different water qualities for different usages (at home, industry, etc.) through secure monitoring sys-tems, always guaranteeing the water quality (especially freshwater) • Advanced multiparameter sensors supporting new capa-bilities, such as stability, antifouling, accuracy, etc. • Real-time microbial- detection and removal are feasible

		• Large-scale deployment of multiparameter devices allowing advanced data analysis in water distribution networks for more intelligent water management • Freshwater quality prediction based on digital twin technology capabilities considering real-time environmental conditions
	Topic 4.2: resource management	• High-performance monitoring systems to identify and quantify the presence of emerging pollutants and high-risk chemical species derived from human action • Integrated vision for all aspects related to water in systemic and non-cyclical areas; process reengineering and redesign of monitoring, control and exploitation systems based on advanced tools for decision-making through the generation of models • Paradigm shift in the vision of water as a cycle to a system that must be optimised
Major Challenge 5: biodiversity restoration for ecosystems resilience, conservation and preservation	**Topic 5.1**: biodiversity restoration for agriculture ecosystem	• Reduction of European cumulated carbon and cropland footprint by 20% over the next 20 years, while improving climatic resilience of European agriculture and stopping biodiversity erosion
	Topic 5.2: biodiversity restoration for aquaculture ecosystem	• Precision aquaculture systems for optimal feeding (minimising waste and feed residuals), optimal use of antibiotics/hormones and optimal use of freshwater
	Topic 5.3: biodiversity restoration for fisheries ecosystem	• Technologies to make fishing gear more selective and environmentally respectful
	Topic 5.4: biodiversity restoration for forestry ecosystem	• Preserve the protected and restored forestry areas, as well as continuing to restore the remaining degraded forests

3.6

ECS Key Application Areas

DIGITAL SOCIETY

3.6. DIGITAL SOCIETY

3.6.1. SCOPE

Supporting the digital transformation throughout society
This chapter describes the type of digital innovations that are essential to stimulate an inclusive and healthy society, and which will in turn contribute to solutions for European challenges in the fields of sustainability health, well-being, mobility, security, energy, and consequently to European economic prosperity.

Europe needs digital solutions that support the individual, and at the collective level to empower society as a whole. These (smart) digital solutions will be massively driven by recent technologies such as Artificial Intelligence (AI), robotics, 6G, virtual reality (VR) and augmented reality (AR),possibly brain–computer interfaces (BCIs). In particular, the recent spectacular advent of Generative AI boosted the application of AI technology and its use for society. These technologies are now shaping new ways of how people use and interact using technological solutions, with each other, and with society and the environment. Digital innovations should facilitate individual self-fulfilment, empowerment and resilience, collective "inclusion" and safety, as well as a supportive infrastructure and environment. However, these innovative technologies, especially AI, also bring risks. AI can bring many advantages, but also lack of transparency, bias and discrimination, privacy concerns, ethical dilemmas, security risks, unwanted dependence, job displacement, and misinformation and manipulation, to name the most important[1].

Furthermore, such transformations are also introducing a wide range of ethical considerations, as digital innovations need to address societal concerns in a sustainable way, guaranteeing participation and reducing inequality. A human-centred approach is therefore a key aspect of the EU's approach to technology development. It is part of European social and ethical values, (social) inclusiveness, and the creation of sustainable, high-quality jobs through social innovation.

In summary, we see four areas to be addressed as our Major Challenges: facilitating individual self-fulfilment, facilitating empowerment and resilience, facilitating inclusion and collective safety, and facilitating supportive infrastructures and sustainable
environments.

3.6.2. APPLICATION TRENDS AND SOCIETAL BENEFITS

3.6.2.1. External requirements

To guarantee economic and societal growth in Europe, digital inclusion and transformation requires tools and infrastructures in application domain roadmaps as described in the other

[1] The 15 Biggest Risks of Artificial Intelligence (forbes.com)

chapters. Digital technology is permeating every aspect of society and is an important instrument of change.

People's expectations of the future impact of technology are broadly positive, but also involve specific concerns around employment, income, safety, equality, and trust. The impact of science and technological innovation on prosperity, individual well-being, sustainability, fairness, and trust is continuously growing, which underlines the importance of investing in our digital strategy today.

This can be further illustrated by the spectacular advent of AI-technology (see figure 3.6.1): the AI market has seen significant growth, offering numerous opportunities for new players. AI, which encompasses technologies capable of tasks requiring human intelligence, is revolutionizing sectors such as healthcare, finance, transportation, and customer service. The market's expansion is largely driven by the increasing demand for automation and efficiency, as businesses adopt AI to streamline operations, reduce costs, and enhance productivity. Additionally, advancements in machine learning and the abundance of big data have further fuelled AI's growth, enabling systems to improve over time and extract valuable insights for better decision-making.

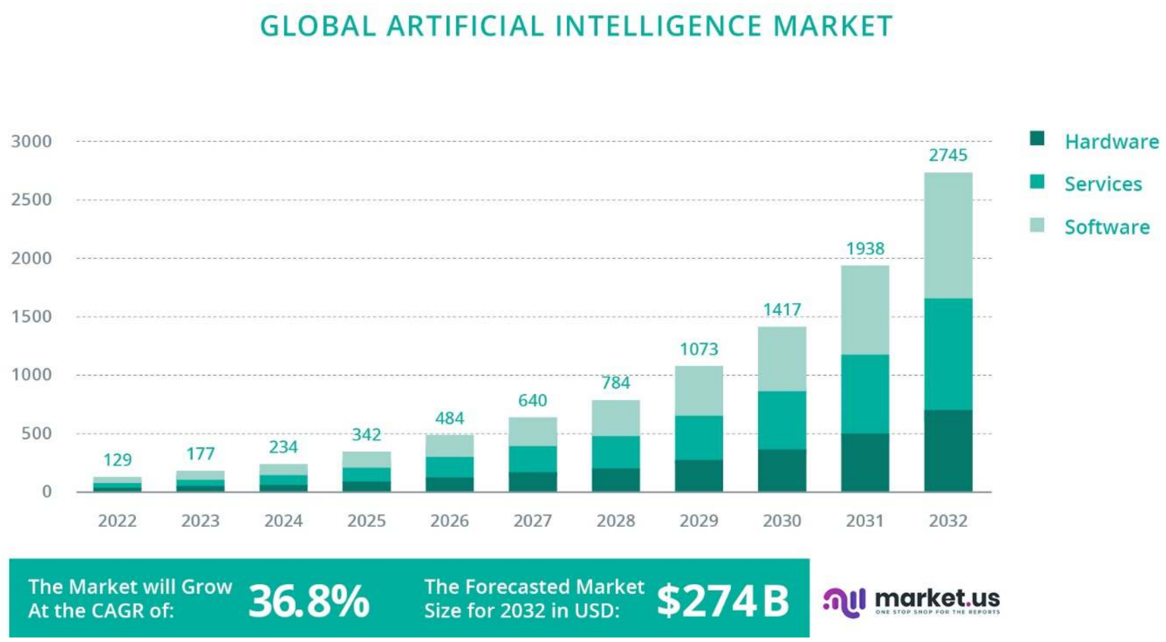


Figure 3.6.1 Growth of global AI market
(source https://market.us/report/artificial-intelligence-market/)

In striving to guarantee European sovereignty to support European digital societal goals, safety, equality, and trust are key requirements. But what does this mean for electronic components and systems (ECS) for our society? Ubiquitous connectivity ("*everywhere and always on*"), online services and social media ("*always online*") drive people to rely on intelligent applications and the services they offer. Public and private infrastructures will increasingly be connected, observed, and controlled via digital infrastructures ("*continuous monitoring*").

Two important further drivers for European society and economy – from a human-centred approach on AI perspective – are lifelong learning and training, as well as being able to work anywhere, any-place. The trend to work from home whenever possible (earlier triggered by the Covid pandemic) has continued, and people are combining work and private life in a better way. In rural areas, as well as in cities, it should be easy to work either from home or remotely in distributed groups/workforces. This can be facilitated through living labs and learning factories at both a personal and collective level.

3.6.2.2. Societal benefits

All of European society will benefit from a major (AI-based) evolution in intelligent systems, on both the individual and collective levels:

- The benefit of digital inclusion for all individuals will involve employability through lifelong learning and training, and the personal well-being of individuals. To achieve these, the key ambition is to maximise the individual development of citizens.
  - How? By ensuring personal resilience, enabling lifelong learning and development, and stimulating employability. Human-centred solutions will optimise services to the needs and capacity of each individual, for applications in areas such as healthcare, lifestyle, coaching, training, and working from home or remotely collaborating in a "distributed" workforce. This will boost employee productivity, improve their work/life balance, and foster better mental health, and reduce pollution from commuters.
- The overall individual benefit is "well-being". A factor such as "prosperity" means job security, material living standards and the right to have the optimum education, any time, any place. On an individual level, well-being means health for every one of every age, and also adequate housing, ensured safety, protected privacy, reliable and ubiquitous digital infrastructures, in addition to social connectedness and more intense social cohesion. Our key objective is to empower and protect the individual.
  - How? By ensuring acceptable and trustable AI technologies to increase inclusion and prevent exclusion, protecting citizens against identity theft, and providing a protective environment against new virus infections; in addition, through lifestyle monitoring and coaching, to enable and support healthier lifestyles.

- The benefits of digital inclusion. A society resilient against setbacks, and the societal acceptance of novel technologies will achieve the key ambition of safeguarding a collective society and well-being for all.
  - How? By societal and digital inclusion, providing societal access for all, and ensuring collective resilience against setbacks. Also, the elderly will be supported to continue their social participation, which will reduce feelings of loneliness, improve their well-being and health, but provide reassurance that their precious experience can still be used.

- On an environmental level, the benefits are a physical and digital sustainable environment, intelligent Infrastructure management, stability and resilience against threats, and agreement on fall-back solutions in times of crises. The main aim is to contribute to a supportive infrastructure and environment.

   - How? By providing reliable and resilient infrastructures, protecting society against destabilising forces, establishing a sustainable environment, and securing controlled climate change. Monitoring and intelligent control of infrastructures will also contribute to a sustainable environment by solutions that address, for example, optimal use of natural resources, reduction of pollution and crisis management.

"Sustainability" in a wide sense implies both environmental and economic sustainability, as well as equal opportunities for all people. It is related to fairness and trust in our societies. It must be ensured that AI-based systems will take European-style human values into account by design (to which the AI-act is contributing). Continuation of a human-centred approach is therefore a key requirement. As such, "FAIRness" (findability, accessibility, interoperability, and re-use) will help to shape future applications too.

3.6.3. MAJOR CHALLENGES

Enabling and ensuring a digital society implies various aspects to be facilitated by trustable ECS products and services. To structure these aspects, we distinguish between the individual or collective context and the internal or external scope. This leads to the matrix shown in 3.6.2 below.

Each of these four areas relates to one of the following Major Challenges:

- **Major Challenge 1**: Facilitate individual self-fulfilment.
- **Major Challenge 2**: Facilitate empowerment and resilience.
- **Major Challenge 3**: Facilitate inclusion and collective safety.
- **Major Challenge 4**: Facilitate supportive infrastructures and sustainable environments.

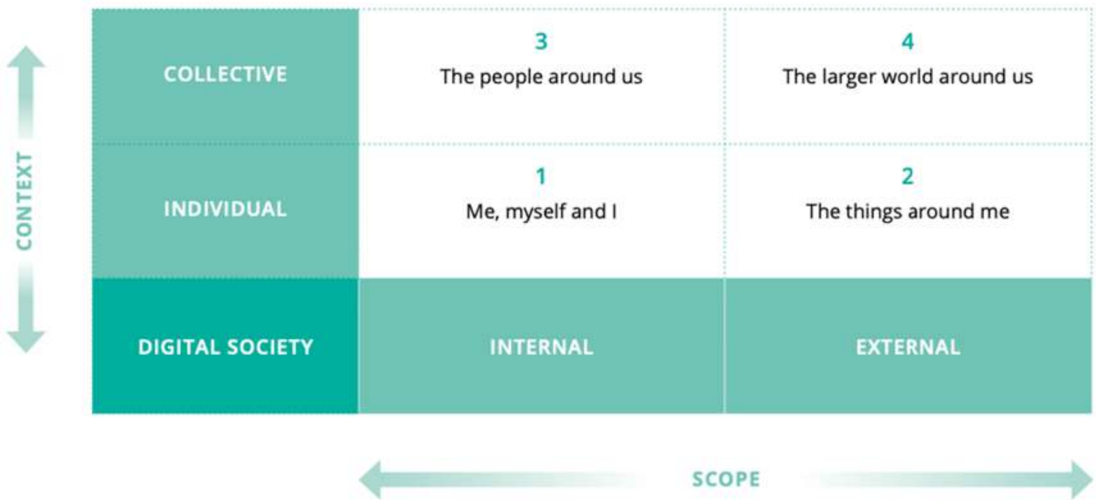


Figure 3.6.2 - Structuring the Major Challenges in scope and context

3.6.3.1. Major Challenge 1: Facilitate individual self-fulfillment

3.6.3.1.1. Status, vision and expected outcome

Ambition: to maximise the individual development of citizens.

- Provide empowerment to citizens and ensure personal resilience.
- Enable lifelong learning and development for both children and adults (serious gaming, including AI-based AR/VR).
- Give citizen more freedom to do their work wherever they want/need.
- Stimulate employability (e.g. by gamification[2], tool and means for stimulation efficient remote work).
- Improved human–machine interaction solutions for perception, reasoning, and autonomy, with interaction being adaptive to the user's abilities.

To maximise the empowerment and self-fulfilment of citizens, Europe has to strive for lifelong learning, employability, and the freedom to work wherever one resides, as well as optimal well-being in the context of an independent and pluralistic media. These enable lifelong empowerment by keeping citizens informed and facilitate the flow of educational content. Educating through the media is an important means to develop valuable skills that will help to end violence and eradicate forms of discrimination (such as sexism and racism). More fundamentally, the media encourages the acquisition of civic knowledge and facilitates discussion concerning current issues, while at the same time entering new frontiers of engagement using on-demand and interactive paradigms, and in employing AR/ VR technologies backed by trustable 5G/6G connectivity.

The 30-year career has become a thing of the past. Education does not end after school; individuals need to keep on learning throughout their careers to stay up to date and adapt their skills as the world changes at an unprecedented rate. To better support lifelong learning, technologies are needed that encourage collaboration, foster autonomy, and responsibility, and implement learning initiatives. Technological advancements such as cloud computing, mobile devices and innovative web technologies are still relatively new additions to the workplace that must be further explored[3].

To provide the citizen with more freedom to do their work wherever they want or need, Europe must ensure the availability of trustable high-bandwidth secure connections (wired and wireless) at all possible locations one could use to work from. This should be reinforced by easy and secure access to cloud applications, and novel AI-based solutions to automate processes, analyse data, guide the user in decision-making, and to minimise repetitive work.

Advanced technologies, including smart automation and AI, have the potential to not only raise productivity and GDP growth, but also to improve well-being more broadly, as well as offer a healthier life and longevity, and greater leisure time. Studies have shown that, besides income, the following factors contribute to individuals' well-being and self-reported life satisfaction: social life, use of leisure, health, spouse/partner, job, flat/house, and the amount

[2] Gamification: The application of game design elements and principles in a non- game context.
[3] https://www.trainingjournal.com/articles/ opinion/how-promote-lifelong-learning- using-technology

of leisure[4]. Innovative technologies in the digital society can, and will, influence all these factors.

There are ample examples of relevant tools which utilises gamification to offer the lifelong free and open learning of languages and brings massive open online course (MOOC) platforms to the public. Additionally, AI-tools that are based on Large Language Models (such as ChatGPT, Claude and others) are beneficial for the individual development of citizens in several ways (access to knowledge, skill development, critical thinking, etc.). However, no tool can replace the richness of human interactions, creativity and experiences in personal development and interaction.

3.6.3.1.2. Key focus areas

High-priority research and development and innovation (R&D&I) areas:

- Digital inclusion: tools, infrastructure, training, connectivity.
- Online education and examination: VR/AR training and support.
- Improved human–machine interaction (HCI) solutions.
- Support devices: wearables, robots, cobots, etc.
- Nudging and serious gaming: for personal development and healthier lifestyles.
- The trusted element: how to be sure that not more is done (under the surface)

To improve the awareness of our body’s condition to external or internal stimuli, smart systems can provide support for disabilities, or a personal coach and trainer to identify behaviour to be avoided (wrong body position and other unhealthy habits), as well as possible future injuries or disorders. Smart systems can also offer an immersive experience through vision, gaming, and sensory interaction by way of VR or AR. Consumers can be offered the immediacy, individualisation, interactivity, and immersion they expect from media content consumption (“even better than being there”).

A healthier and more comfortable environment can be offered based on personal preferences (control of temperature, humidity, air flux, etc.), in the context of running activities and clothing, and by adapting lighting and acoustic quality to one’s own sense of well-being. It also provides the capability to comfortably communicate and interact remotely with people, institutions, and sellers, possibly without leaving home, saving time for self- development and leisure.

Selective automation, AR at work and a range of feedback tools can help boost satisfaction and give more meaning to work. This is a particularly important element for the millennial generation, which -according to surveys- tends to place more emphasis on work satisfaction than on income (above a certain income level). Technological advances have made it possible to place audiences in the middle of the action and to offer them immediacy, individualisation, interaction, and immersion without the need for them to actually be there in person. This will further change consumption patterns and create new business opportunities.

---

[4] https://www.mckinsey.com/featured- insights/future-of-work/tech-for-good- using-technology-to-smooth-disruption- and-improve-well-being

Furthermore, we have to mitigate the risks of AI and other functionality that is based on massive data storage in the cloud. This can be done in different ways, such as:

- Introduce fake news detection by analysing AI-generated content on social media
- Develop technologies to prove the source of content creation, for instance by showing that a video was created with an actual camera. If part of the processing is done externally and/or in the cloud, add traceability of what was done where.
- Add watermarking to created content to proof that it is not artificial.
- Secure control of systems used for content creation, for instance by prevention of production equipment to be taken over by external aggressors in big events like Olympic Games, Eurovision song festival, European and World championships, etc.

Required R&D&I developments within ECS

Taking the above into account, specific R&D developments are necessary within ECS technology, as shown in

Specific R&D developments necessary	**ECS technologies**							
Major challenge 1: “Facilitate individual self-fulfilment”	Process techn, equipm, materials & manufact.	Compoinents, miodules and system integration	Embedded SW & Beyond	System of Systems	Edge Comp. & emb. AI	Connectivity	Architecture & Design	QRSC
	1.1	1.2	1.3	1.4	2.1	2.2	2.3	2.4
Reliable, dependable and secure SW and HW	x	x	X	x	x	x	x	X
Mature human systems interaction methods				X			X	X
Trustable AI/Machine Learning algorithms		x			X			X
Energy-efficient HW and SW solutions (e.g. for IoT devices, wearables)	X	x	X		X			
Seamlessly operating SW (e.g. for IoT devices, wearables)		X	X	X		X		
Ubiquitous, reliable, and energy-efficient connectivity	X	X	X			X	X	X

Figure 3.6.3.

Specific R&D developments necessary	**ECS technologies**							
Major challenge 1: “Facilitate individual self-fulfilment”	Process techn, equipm, materials & manufact.	Compoinents, miodules and system integration	Embedded SW & Beyond	System of Systems	Edge Comp. & emb. AI	Connectivity	Architecture & Design	QRSC
	1.1	1.2	1.3	1.4	2.1	2.2	2.3	2.4
Reliable, dependable and secure SW and HW	x	x	X	x	x	x	x	X
Mature human systems interaction methods				X			X	X
Trustable AI/Machine Learning algorithms		x			X			X

Energy-efficient HW and SW solutions (e.g. for IoT devices, wearables)	X	x	X		X			
Seamlessly operating SW (e.g. for IoT devices, wearables)		X	X	X		X		
Ubiquitous, reliable, and energy-efficient connectivity	X	X	X			X	X	X

Figure 3.6.3 - Required R&D&I developments within ECS –Major Challenge 1

3.6.3.2. Major Challenge 2: Facilitate empowerment and resilience

3.6.3.2.1. Status, vision and expected outcome

Ambition: empower and protect the individual citizen.

- Increase inclusion and prevent exclusion.
- Protect citizens against cyber-fraud (scams) and identity thefts; provide privacy.
- Enable safe smart homes with ubiquitous connectivity.
- Ensure acceptable AI technologies.

Diversity and inclusion within societies are increasingly recognised as crucial for equality at work and economic development. Research has established a strong link between gender equality in society, attitudes, and beliefs about the role of women, and gender equality in work[5]. Technology can improve equality at work – for instance, by revealing pay gaps and biases, and helping de-bias recruitment. It can also improve equal access to essential services – for example, biometrics and cloud technology can contribute to increasing the diffusion of microfinance to women and underserved populations. Technology can also help enforce inclusive legal rights, policies, and social norms. While e-voting still poses a number of cybersecurity challenges, it can support diversity by facilitating the vote for vulnerable and marginalised parts of society. Finally, technology can help with physical security and autonomy for minority groups through objects and digital communications tools that reduce or mitigate exposure to risk – for example, connected devices such as smart bracelets can enable women to signal an assault and call for help.

Reliance on technology comes with many benefits, but also brings new risks[6]. The radical nature of the ongoing technology transition could result in risks that are not just an extension of the previous challenges but require fundamental changes to core aspects of our society, including how we think about our identity, security, and rights. Concerns about technology are justified by recent events, such as security breaches in prominent companies, data theft and information misuse. In addition, AI provides more powerful examples of potential risks. Its full potential can be used only if we fully rely on it for decision-making, allowing it to process data beyond the human ability to cross-check and verify. This depends on a high level of trust, raising questions about, and requiring, new technical solutions that take into account explainability, accountability, trustworthiness and ethics.

---

[5] McKinsey Global Institute: “Tech for Good: Smoothing disruption, improving well- being”, May 2019, p 42 and p 43.

[6] McKinsey Global Institute: “Tech for Good: Smoothing disruption, improving well- being”, May 2019, p 58.

In the early ‘20’s we have experienced the necessity of a connected smart home and an adequate home office during the pandemic. However, the availability of high bandwidth connectivity is not evenly distributed geographically across Europe.

Machine learning is essential for a resilient future. AI will have a far greater chance of successful implementation if there is a focus on four key areas: augmented intelligence; intelligent automation; assessed intelligence; and adaptive intelligence. Augmented intelligence concerns augmenting and thus improving the productivity of humans. Intelligent automation is about building systems that integrate humans and machines in productive ways (instead of just replacing humans entirely with machines). Assessed intelligence is all about making models robust by evaluating them rigorously and continuously. Finally, adaptive intelligence involves developing more resilient systems that can adapt to changing circumstances by shifting to a causal inference paradigm.

3.6.3.2.2. Key focus areas

High-priority R&D&I areas:

- Reliable and ubiquitous digital infrastructures.
- Access control/intrusion detection/surveillance.
- Provide protective environment and tools against virus infections.
- Protect individual citizens against cyber-fraud (scams) and identity theft.
- Off-grid living and emergency survival.

Since the Covid pandemic, working from home has become an integral part of how knowledge workers do their work. To further enable working from home (or wherever and whenever one wants), wireless and wired infrastructures will have to be further improved (through increased reliable bandwidth, lower cost, better geographical coverage and finer granularity), security of connections will have to improve to protect the worker at home (as will the company using a distributed workforce with many internet connections) against cyber-attacks, and the theft of personal and/or company information. New functionality running in the private/public cloud will be needed to support real-time actions that may suffer from latency issues over the internet, as well as to support the worker in decision-making. Examples here are control of robotic surgical devices, remote control of robots in industrial processes, remote control of cameras in security applications and live television productions. Other professions, such as translation services, voice recognition and all kind of analytical algorithms for data analysis, also come to mind.

To create equal opportunities, innovative research should include: speech-generating devices (SGD) to help people with speech disorders; exoskeletons that empower disabled people in their everyday life; semi-autonomous vehicles that increase mobility for people with deafness and blindness; smart objects linked to geospatial information to improve women’s security (e.g. invisible SOS buttons); augmentative and alternative communication tablets that help paralysed patients; VR solutions that provide realistic experiences for people with physical disabilities; and smart glasses that can be used to help people with autism on cognitive, social and emotional skills.

Given the experience of the past pandemic, we learned that Europe needs better technologies: (i) to fight and contain the rapid spread of highly contiguous diseases (such as Covid-19); and (ii) to ensure that public health institutions can maintain their capacity to meet the ever-increasing needs caused by such a pandemic[7]. The in-depth analysis provided by the European Parliamentary Research Service's "Ten Technologies to Fight Coronavirus" identifies the importance of AI, blockchain, open-source-, telehealth- and gene-editing technologies, 3D printing, nanotechnology, synthetic biology, and drones and robots for fighting pandemics.

Intrinsically, technology is neither good nor bad – it is the use to which it is put that makes the difference. Malicious uses of technology include mass disinformation campaigns and cyber-attacks that seek to jeopardise national security, and cyber-fraud that targets consumers. This duality has always existed. Over the coming years, technologies such as the IoT, smart robotics, automation and AI are likely to follow the same pattern. It is up to European technology specialists to ensure that the technologies developed not only support diversity and inclusion, but also protect both the individual and groups against cyber-attacks, theft of personal information and unwanted intrusion into the personal environment.

Required R&D&I developments within ECS

To facilitate empowerment and resilience, specific R&D developments are necessary within ECS technology, as shown in

Specific R&D developments necessary	**ECS technologies**							
Major challenge 2: "Facilitate empowerment and resilience"	Process techn, equipm, materials & manufact.	Compoinents, miodules and system integration	Embedded SW & Beyond	System of Systems	Edge Comp. & emb. AI	Connectivity	Architecture & Design	QRSC
	1.1	1.2	1.3	1.4	2.1	2.2	2.3	2.4
Reliable, dependable and secure SW and HW	X		X			X	X	X
Trustable AI/Machine Learning algorithms					X			X
Advanced cyber-security and privacy methods and tools		x	x			X	X	X
Ensuring of safety and resilience based on ECS technologies		X	X	X	X	X	X	X
Energy-efficient and dependable HW and SW solutions (e.g. for IoT devices, wearables)	X		X	X	X		X	X
Seamlessly operating SW (e.g. for IoT devices, wearables)		X	X	X		X		
Ubiquitous, reliable, and energy-efficient connectivity and localization	X	X	X			X	X	X
Secure broadband connectivity based on 5G systems and beyond	X	X	X			X		X
Distributed (production) systems		X	x	X	X	X	X	X

[7] https://www.europarl.europa.eu/ RegData/etudes/IDAN/2020/641543/ EPRS_IDA(2020)641543_EN.pdf

Figure 3.6.4.

Specific R&D developments necessary	**ECS technologies**							
Major challenge 2: "Facilitate empowerment and resilience"	Process techn, equipm, materials & manufact.	Compoinents, miodules and system integration	Embedded SW & Beyond	System of Systems	Edge Comp. & emb. AI	Connectivity	Architecture & Design	QRSC
	1.1	1.2	1.3	1.4	2.1	2.2	2.3	2.4
Reliable, dependable and secure SW and HW	X		X			X	X	X
Trustable AI/Machine Learning algorithms					X			X
Advanced cyber-security and privacy methods and tools		x	x			X	X	X
Ensuring of safety and resilience based on ECS technologies		X	X	X	X	X	X	X
Energy-efficient and dependable HW and SW solutions (e.g. for IoT devices, wearables)	X		X	X	X		X	X
Seamlessly operating SW (e.g. for IoT devices, wearables)		X	X	X		X		
Ubiquitous, reliable, and energy-efficient connectivity and localization	X	X	X			X	X	X
Secure broadband connectivity based on 5G systems and beyond	X	X	X			X		X
Distributed (production) systems		X	x	X	X	X	X	X

Figure 3.6.4 - Required R&D&I developments within ECS – Major Challenge 2

3.6.3.3. Major Challenge 3: Facilitate inclusion and collective safety

3.6.3.3.1. Status, vision and expected outcome

Ambition: safeguard collective society and well-being for all.

- Societal and digital inclusion.
- Provide societal access for all.
- Ensure collective resilience against setbacks.

Although European countries have diverse types of welfare models, they also share a history of robust social protection and a focus on inclusive growth, which has been under stress in recent years[8]. There could be cracks in the sustainability of the EU social contract over the next decade caused by six trends: ageing; digital technology, automation, and AI; increased global competition; migration; climate change and pollution; and shifting geopolitics. Based on these trends, inequality may rise again, and divergence within Europe could increase.

Inequality at work may emerge through a combination of: (i) automation and the substitution of labour; and (ii) corporate diffusion dynamics, leading to a competitive disadvantage among

[8] Testing the resilience of Europe's inclusive growth model, McKinsey Global Institute, December 2018, p4.

non-adopting firms. To prevent reduced employment and secure real wage growth, automation using AI, robotics and other innovative technologies should lead to significant productivity gains. In general, occupations based on more repetitive and non-digital tasks will be taken by workers with low education and skills, who will therefore be the first to experience pressure on wages.

Collective growth and well-being are not only determined by equality at work, but also by individual development supported by collective interactions. Studies have shown that active social relationships increase health and longevity by improving key biomarkers of physical health. A lack of interaction causes a subtle decline in mental health by reducing attention, learning, memory, and decision-making skills. In short, our bodies reward us for social interaction and punish us for isolation by negatively impacting mental and physical health. Direct interactions with family and friends, participating in team sports and, for instance, visiting an event with friends are particularly important. No technology can fully replace direct contact. However, during the former Covid pandemic, we had to rethink our social interactions, and to adapt technologies to increase and improve social interactions – not just between individuals, but also between individuals and groups, as well as between different groups. How can improved technologies support existing social interactions, and secure healthy digital social interactions in cases of setbacks?

It may seem obvious that the Electronic Components and Systems used are trustable devices and technology, but this is an important boundary condition in Europe.
In addition, collective safety can be enhanced by solutions that directly address specific communities or groups of people the individual is a member of, such as family, friends, neighbourhood, region, (sports)club or association. These solutions can either have a warning or alerting function (e.g. contamination, local fire, local air pollution, incident of violence), but can also be of a supporting nature – for instance, alerts or instructions in combination with collective supporting devices (e.g. automated external defibrillator (AED), diagnostics, measurement).

3.6.3.3.2. Key focus areas

High-priority R&D&I areas:

- Digital inclusion: tools, infrastructure, training, connectivity.
- Collective safety: secure access control, surveillance, pandemic control, prevention of misinformation without limiting freedom of expression.
- Safe environment for living, working and transport: buildings and bridges resilient against earthquakes through continuous monitoring (e.g. fibre-based stress sensors).
- Emergency/crisis response solutions and services.
- Dynamics of society: systemic change.

As Europe wants to play a key role in digital inclusiveness, it is important to ensure availability and accessibility of solutions to enable remote education, learning, training and assessment of professionals, students, and consumers in all regions (both cities and rural areas). Also, solutions to support social inclusiveness for people of all age should become available.

The EU has stated, in their document on orientations towards the first strategic plan for implementing Horizon Europe[9], that the interaction of science, technology, social sciences and humanities will be crucial in this respect, as will be the input of the creative sector and artists to sustainable inclusive innovation and human- oriented technologies.

To facilitate inclusion, more research will be needed on education, simple human–machine interfaces and digital technology interfaces that avoid the digital split between high- and low-educated citizens. In addition, remote presence, and remote connectivity to keep people connected even if they are not in the same location, trustworthy social media, serious gaming, media consumption and AR/VR will be key.

To safeguard digital inclusion, education is one of the most important research areas. Examples here are the use of AI to build personalised journeys and enhance learning outcomes, to adapt curriculum to individual student needs, digital support and nudging systems to reduce the administrative burden on teachers, tablet-based learning to improve results and decrease distress for students with dyslexia, automation of administrative tasks to free up time and resources for educational professionals, wearable devices that provide real-time support to pupils, eye-tracking solutions to adapt students' learning experiences, and use of AR/VR to provide immersive experiences to civilians in less well-served areas.

AR may improve connectedness for remote places, reducing the need for commuting or business travel. It could also enable consumers to enjoy an event together even if they are not physically at the event.

There are still several challenges to effectively take full advantage of AI in video creation and consumption. One is the size of video data. Results are only accurate when algorithms are fed with millions of observations. Technologies therefore have to be deployed, and strategies have to be implemented to gather data at scale to harness the full power of AI techniques. However, size creates another challenge: datasets need to be manually labelled by humans to train the model, making the process expensive and cumbersome. Many new techniques are becoming available to overcome the challenge of data categorisation, such as reinforcement learning, generative adversarial networks, transfer learning and "one- shot learning" and large-language-models. In consumer-facing applications, such as marketing and recommendation algorithms, AI models may need to be refreshed continuously due to changes in the environment that drives them. Continuous updates to AI models are expensive. Other challenges relate to data management and data gathering: to create accurate results with AI, and thus value, diverse types of data have to be managed in a unified manner. This includes audience data, operational data, and content data (metadata). Also, "selection bias" (i.e. the data gathered is not representative of the population studied) has to be prevented to exclude wrong conclusions in a perfectly working model.

To facilitate collective safety, further research is required on secure access control, intrusion detection, (video) surveillance of security sensitive areas, and individual and collective activity tracking.

---

[9] https://ec.europa.eu/research/pdf/horizon- europe/ec_rtd_orientations-towards-the- strategic-planning.pdf

Secure access control as a service (ACaaS) is growing in relevance. This combines biometric readers and identity access management and can be integrated with other physical security systems (e.g. video surveillance) and building automation systems. Combined with building occupancy management systems, it can deliver valuable information on the location of staff and visitors, and in the event of an emergency to rapidly clear the building.

Both the former Covid pandemic and the military conflicts on our continent has brought new physical security requirements. In addition to regular cameras, thermal cameras as well as different sensing devices could be added at the entrance of buildings and venues to measure people's state (e.g. temperature) and investigate if they do not carry dangerous materials as they enter premises. Physical access control, enriched with video security evidence, can provide important insights on where an infected individual has been, which doors they have used and who else may have come into contact with those doors and that individual. It can also provide these insights for more general security purposes.

More research on AI security solutions will ease the work of security operators. AI software can analyse images and audio from video surveillance live streams and recordings, and use image recognition algorithms to recognise faces, objects, and events, more than a hundred times faster than human operators. AI algorithms can also be used to carry out event detection, scene reconstruction, video tracking, object recognition, and (re)-identification, 3D pose estimation, motion estimation and image restoration. Video surveillance may be extended with freely moving cameras mounted under drones to recognise unusual behaviour in crowds from a high altitude, to monitor hazards such as fires, floods, or erupting volcanoes, and to recognise criminal faces and follow targets. Since drones are airborne, they need fast mobile and wireless communications. Low-latency broadband technologies such as 5G/6G can improve the precision and speed of their response times and enable high-speed communication to a nearby edge computing device.

Video quality should be further improved to support deep-learning algorithms, and to improve the video experience in media consumption: the spectral range and colour gamut can be extended, sensitivity must increase for low light use and especially dynamic range for better performance under all (and changing) lighting conditions as this has the greatest impact on the perception.

AI video and audio algorithms will have to be transparent and explainable. Dedicated video and audio technologies will be required to prevent and trace fake video and audio used to create misinformation in (social) media. Audio and video equipment used to create content must be able to watermark content streams to prove authenticity, and metadata should be added to the streams to prove what processing activities have taken place from image capture until display at home. Next to that equipment used for content creation based on open public infrastructures (such as internet and 5G) and with processing or control activities in the cloud must be better secured to prevent unwanted take-over by aggressors.

Required R&D&I developments within ECS

To facilitate inclusion and collective safety, specific R&D developments within ECS technology are necessary, as shown in Figure 3.6..

Specific R&D developments necessary	ECS technologies							
Major challenge 3: "Facilitate inclusion and collective safety"	Process techn, equipm, materials & manufact.	Compoinents, miodules and system integration	Embedded SW & Beyond	System of Systems	Edge Comp. & emb. AI	Connectivity	Architecture & Design	QRSC
	1.1	1.2	1.3	1.4	2.1	2.2	2.3	2.4
ECS technologies for AR/VR and high-quality video/videoconferencing	X	X	X		X	X	X	
Tools, methods, SW and HW technologies for extensive and ubiquitous use of AI/Machine Learning	X		X		X	X	X	
Advanced cyber-security and privacy methods and tools						X	X	X
Intelligent connected IoT devices using new sensors for safety and resilience of EU societies	X	X	X	X	X	X	X	X
Ubiquitous, reliable, and energy-efficient connectivity and localization	X	X	X			X	X	X
Secure broadband connectivity based on 5G systems and beyond	X	X	X			X		X

Figure 3.6.5 - Required R&D&I developments within ECS – Major Challenge 3

3.6.3.4. Major Challenge 4: Facilitate supportive infrastructure and a sustainable environment

3.6.3.4.1. Status, vision and expected outcome

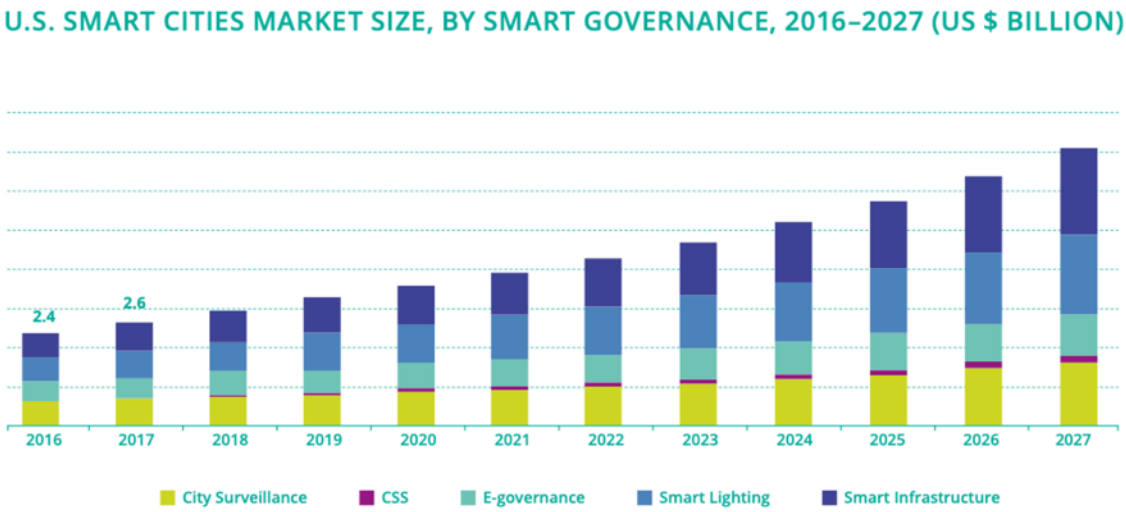
Ambition: contribute to a collective supportive infrastructure and environment.

- Provide trustworthy, reliable, and resilient infrastructure.
- Establish a sustainable and secure environment.
- Provide means for controlled climate change.

To fully benefit from the power of digitisation, Europe must enable a supportive infrastructure and environment. Given the rapid pace of change, this requires companies to get their technology, people, and culture ready to join the digital transformation. This should be achieved by providing a reliable and resilient digital infrastructure (with ubiquitous and continuous connectivity), protecting society against destabilising forces and establishing a secure and sustainable environment. The former includes preventing harmful use of the internet (e.g. manipulation of elections, misinformation such as "deepfakes" and "cheapfakes", but also identity theft and phishing), which are covered by Major Challenge 3. The latter includes securing controlled climate change (as stipulated in the Green Deal) by providing adequate means that can be utilized by citizens. Furthermore, monitoring, and intelligent control of infrastructures, essential resources, and their recycle process (especially in the urban environment) will contribute to a sustainable environment.

The vision is to introduce new digital products and services that contribute to a sustainable lifestyle in all areas of human life, including cradle-to-cradle and circular economy aspects. We are addressing the following aspects:

- Comprehensive assessment of resource usage to identify largest areas of energy and materials consumption and CO_2 production. As well as air quality monitoring systems, they need to offer solutions for lighting, heating, computing with reduced usage of energy, and other resources. In addition, solar panels and batteries, home-grown vegetables and city farming systems are key.
- Providing smart systems based on IoT and unmanned/robotic platforms that provide secure environments and support digital business life with the minimum amount of resources (energy, water, paper, travelling, etc.), ensuring a highly efficient, productive, and sustainable working environment.
- Smart water management to protect resources. Intelligent management of energy in public spaces such as football stadiums and railway stations, including smart street lighting. Promoting green areas in cities and enabling citizens to provide their own sustainable solutions. Reduction of (food) waste in supermarkets and restaurants, as well as resource recycling.
- Development of new digital solutions that can be used by citizens themselves or be used to stimulate societal attitudes and activities towards sustainability and addressing climate change, e.g. car sharing, use of public transportation, bikes, and e-scooters.



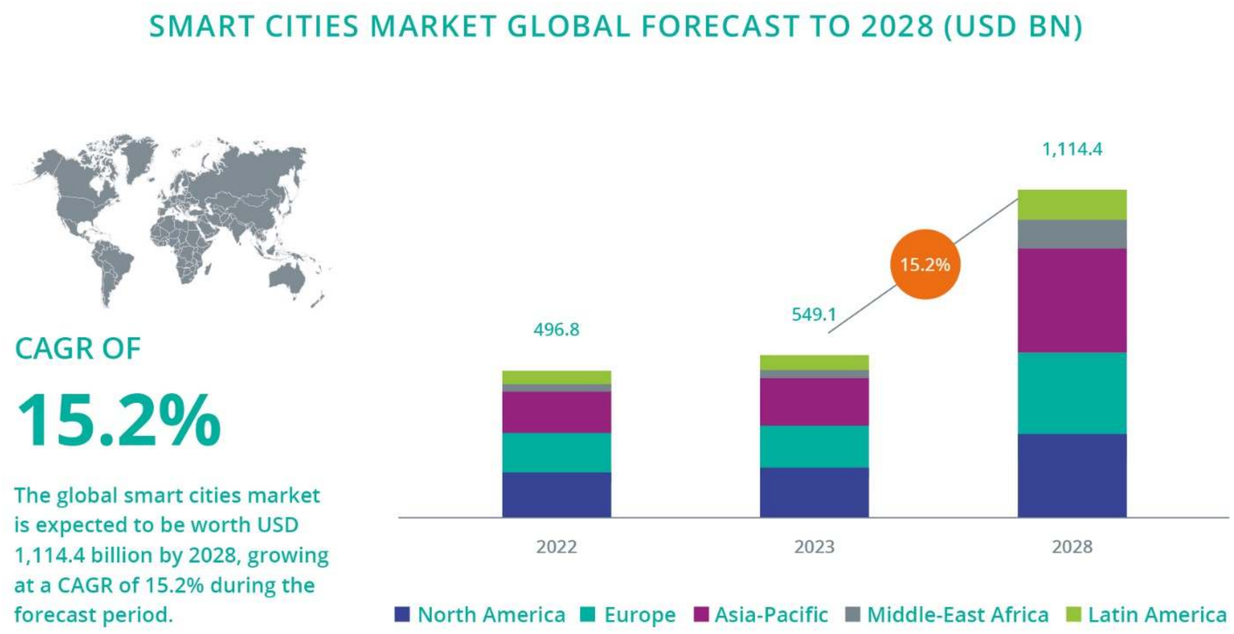


Figure 3.6.6 - Growth of smart cities market (Source: www.marketsandmarkets.com)[10]

The European approach to working with regard to digitalisation will be focused on the preservation of our democratic system, and on values such as trust and cooperation. Ethical requirements will include fairness, accuracy, confidentiality, transparency, accountability, explainability, trustworthiness and absence of bias. This involves offering AI capability maturity programs to companies that use AI in their designs, to coach them in the best ethical points of view. In this way, products will become more resilient, accessible, reliable, and trustworthy, and hence ready to take part in the new European digital society (with an effective AI act).

3.6.3.4.2. Key focus areas

High-priority R&D&I areas:

- Physical infrastructure management/physical resilience and security.
- Intelligent infrastructure management (intelligent buildings, city-owned infrastructure, synergies with industry, etc.).
- Digital solutions and infrastructure for resources management, digital resilience, sustainability, e-government and citizen support.
- Resource and environment monitoring (air, water, etc.) and feedback to enable more effective management.

To further improve digital transformation towards more secure, effective, and sustainable environments, new digital tools employing IoT, unmanned vehicles/robots and AI/ML-enhanced algorithms have to be infrastructures, investments should be aimed at enhancing infrastructure coverage and quality . Also, outcomes have to be influenced through legal frameworks and by setting standards.

---

[10] Source: Grand view Research: "Smart Cities Market Size, Share & Trends Analysis" Report by Application (Governance, Environmental Solutions, Utilities, Transportation, Healthcare), By Region, And Segment Forecasts, 2020 – 2027. See https://www.grandviewresearch.com/ industry-analysis/smart-cities-market

Intelligent buildings will require security, eco-friendship and building management. Security systems such as access control and cybersecurity were covered under Major Challenge 3, but the further development of smart lighting, air quality monitoring and control, and IoT-based real-time monitoring of electric, water and gas meters to increase the energy efficiency of buildings with the help of distributed energy systems will improve the well-being of occupants and reduce the carbon footprint of buildings. Smart technology (e.g. sensors placed around radiators, boilers, pumps, and other machinery to detect critical levels of noise, vibration, or heat) will enable facility managers to save maintenance costs by switching from a reactive to a predictive maintenance model.

Cities are very complex organisms. They combine a variety of means allowing for mobility, city infrastructure providing diverse types of media (gas, water, energy, etc.), and citizen-oriented services that increase their quality of life. It is predicted that by 2050 between 68% and 90% of the global population might live in cities, from small municipalities right up to megacities11. This means that, in the near future, technical means will be required to enable digital solutions for more sustainable development in cities of all size and wealth. Available technologies from tech giants such as IBM, Microsoft, Amazon, Google, and Cisco raise concerns from city managers about data privacy policies, and the very high maintenance costs caused by licence fees and the potential for vendor lock-ins[12]. Available open-source solutions – such as the Red Hat integration platform, which could be used in smart city applications – can also easily be acquired by large companies such as IBM[13] to be integrated with their company product portfolio offered commercially. This means that, in such a dynamically changing world, open-source solutions that are widely available, promoted and deployed within EU (such as FiWARE[14]) have to be developed to protect European sovereignty and values. Additionally, due to the rich industrial heritage in many EU countries, opportunities for re-using or integrating available well- developed open-source industry platforms, such as the Eclipse Arrowhead Framework[15], can be adapted to smart city needs based on requirements gathered in EU-funded projects. This is especially the case since industry sites are often integrated within city areas, and therefore naturally create synergies that can influence each other. These smart city applications create natural synergies with the System of Systems, Mobility and Digital Industry sections.

The impact of technology on environmental sustainability is likely to be highly significant. In retail, where shifting customer habits will be key (for example, for new products such as plant- or insect-based food), IoT sensors and devices will also yield a positive impact – for example, by reducing waste through improved food temperature or expiry date management. In the manufacturing sector, smart building applications related to energy and wastewater management, as well as applications such as carbon capture and biofuel generation on industrial sites, will have a significant impact.

---

[11] https://unhabitat.org/wcr/

[12] https://www.smartcitiesworld.net/news/ news/city-governments-fear-vendor-lock-in- from-iot-platforms-3776

[13] https://www.networkworld.com/article/967721/ibm-fuses-its-software-with-red-hats-to-launch-hybrid-cloud-juggernaut.html

[14] https://www.fiware.org

[15] https://www.arrowhead.eu

Required R&D&I developments within ECS

Development of supportive infrastructure and a sustainable environment within EU needs the following specific R&D developments within ECS technology:

Specific R&D developments necessary	**ECS technologies**							
Major challenge 4: “Facilitate supportive infrastructure and a sustainable environment”	Process techn, equipm, materials & manufact.	Compoinents, miodules and system integration	Embedded SW & Beyond	System of Systems	Edge Comp. & emb. AI	Connectivity	Architecture & Design	QRSC
	1.1	1.2	1.3	1.4	2.1	2.2	2.3	2.4
Open systems and platforms for managing complex cross-connected physical infrastructure and associated processes		X	X	X			X	
Energy-efficiency oriented HW technologies and embedded SW	X		X		X	X	X	
Advanced cyber-security and privacy methods and tools						X	X	X
Intelligent connected IoT devices using new sensors for safety and resilience of EU societies	X	X	X	X	X	X	X	X
Ubiquitous, reliable, and energy-efficient connectivity and localization	X	X	X			X	X	X
Secure broadband low latency connectivity based on 5G systems and beyond	X	X	X		X	X		X
Distributed (production) systems		X		X	X	X	X	X

Figure 3.6.7 - Required R&D&I developments within ECS – Major Challenge 4

3.6.4. KEY ENABLING METHODOLOGIES

Key Enabling Methodologies[16] support in bridging the opportunities of the Key Enabling Technologies in ECS (electronic components and systems) with the four Major Challenges for the Digital Society. Key Enabling Methodologies (KEMs) are the ‘instruments’ that direct and structure the way of working in multi-collaborative settings, give direction and realise impact with useful applications and meaningful interventions.

KEMs contribute to the integration of the technological opportunities of ECS with the knowledge from design, social sciences, and the humanities about the Digital Society.
Several categories cover the main areas of the KEMs: Vision & Imagination, Participation & Co-creation, Behaviour & Empowerment, Value Creation & Upscaling, Institutional & System Change. KEMs can support in answering questions such as how technological tools,

[16] https://kems-en.clicknl.nl

infrastructure and training can empower citizens towards digital inclusion. How to imagine and anticipate for setbacks and build collective resilience? How to include the relevant societal stakeholders into the R&D&I developments within ECS?

3.6.5. TIMELINE

The following table illustrates the roadmaps for Digital Society.

MAJOR CHALLENGE	TOPIC	SHORT TERM (2025–2029)	MEDIUM TERM (2030–2034)	LONG TERM (2035 and beyond)
Major Challenge 1: facilitate individual self-fulfilment	Topic 1.1: improved human– machine interaction solutions	Intensive research on human–machine interaction solutions	Improved human–machine interaction solutions in commercial phase	Ubiquitous human–machine interaction solutions
	Topic 1.2: online education and examination	Developments of methods and solutions for online education and examination	Online education and examination widely used across the EU	Online education and examination widely used world-wide
	Topic 1.3: VR/AR training and support	VR/AR pilots, including remote training, support, and work	VR/AR training, support, and remote work is mature	VR/AR training widely used across the EU
	Topic 1.4: support devices (wearables, robots, cobots, etc)	Smart watches and robots are commonly used devices	Support devices (wearables, robots, cobots, etc) gain more intelligence and interaction capabilities	Intelligent support devices (wearables, robots, cobots, etc) used in daily life
	Topic 1.5: nudging, gamification (for development or health reasons)	New nudging, gamification systems developed for education and health	Nudging, gamification pilots in education and health	Nudging, gamification (for development or health reasons) is widely used across the EU
Major Challenge 2: facilitate empowerment and resilience	Topic 2.1: access control/ intrusion detection/ surveillance	Developments of new sensors, devices, and algorithms for surveillance systems	Smart multimodal surveillance with AI-based intrusion detection	Smart multimodal surveillance with AI-based intrusion detection used world-wide

	Topic 2.2: reliable and ubiquitous digital infrastructures	Increased quality of service (QoS) and available bandwidth with 5G/6G, less time-critical functions moving to the cloud	Bandwidth and QoS increase especially for video-based applications Time-critical functions moved to cloud or edge-cloud systems	Bandwidth and QoS no longer an issue for video applications. AI algorithms support supervision. Edge-cloud systems are commonly used
	Topic 2.3: social media/serious gaming/AR/VR	AR on social media moves from photos to video >80% on social media in video by 2024; in-game systems that self-adapt to guide human learning. Use of cloud processing solutions.	Apart from, AR also VR for videos on social media Multimodal and multi-sensory interfaces in serious gaming Application beyond single game. Personal learning. Secured IP and cloud connected media content creation.	Real-time emotion state sensing Cognitive learning. Fully transparent , encrypted, and protected content creation with automation in the cloud.
Major Challenge 3: facilitate inclusion and collective safety	Topic 3.1: digital inclusion: tools, infrastructure, training, connectivity	Development of technologies (AR/VR, hearables, haptics, etc) for digital inclusion	Pilot deployments of hybrid systems for collective interactions	Technologies for immersive collective interactions
	Topic 3.2: resilient society against setbacks	Emergency/crisis response solutions and services with ubiquitous localisation	Trustable solutions for collective activity tracking, access control and intrusion detection	Trustable AI-supported hybrid solutions for resilient society
	Topic 3.3: societal acceptance of novel technologies	Technologies (serious gaming, nudging, etc) for societal acceptance and adaptation	Human-oriented trustable AI systems and technologies	Trustable AI for collective growth and well-being
Major Challenge 4: facilitate supportive infrastructure and environment	Topic 4.1: physical infrastructure management/ physical resilience	Development of IoT and dedicated robot-based inspection systems supported by AI algorithms	Pilot deployments of trustable AI-based systems relying on dependable edge/cloud IoT and unmanned solutions	Intelligent, affordable, and trustable IoT, robot-based systems, and unmanned solutions are available

	Topic 4.2: intelligent infrastructure management	Development of systems for intelligent management of infrastructure (water, street lighting, heat, etc)	Pilot deployments of trustable AI-based orchestration systems to create synergies in infrastructure management	Smart systems for multi-domain infrastructure orchestration and management available
	Topic 4.3: digital infrastructure management/digital resilience and cybersecurity	Acceleration of initiatives to create open, secure privacy- oriented systems; development of AI-based algorithms for increased cybersecurity	Adaptation and pilot deployments of available interoperable open and reliable systems supported by trustable AI algorithms for increased cybersecurity	Open, secure, interoperable, and reliable privacy-oriented systems empowered by trustable AI-based IoT solutions available
	Topic 4.4: surveillance, homeland security and emergency response systems	Edge/cloud solutions, IoT systems and robot-based inspection platforms, increased multimodal situational awareness, ubiquitous localisation	Deployment of trustable AI-based edge-cloud solutions, IoT systems and robot-based inspection platforms for surveillance and emergency response support	Trustable and dependable AI-based IoT systems and robot-based inspection platforms for increased situational awareness widely available

4

Strategic Research and Innovation Agenda 2025

LONG-TERM VISION

4. LONG TERM VISION

4.1 INTRODUCTION

In this Chapter, we present research subjects that need to be addressed by European organizations to enable and support effective development of European industry in about a decade from today. The previous chapters have presented status, trends, and plans for the near future, including challenges that are foreseen to require special attention within the coming decade. We build upon these identified challenges and specify long-term industrial needs. These needs are the basis for research programs for effective research and development in appropriate technological and/or application domains, so that European technological strength increases continuously in time and at the appropriate rate. Since lead-time from a first scientific breakthrough (TRL1) to market presence of related products (TRL9) is typically 10 years or more, the effective identification of the future industrial needs is a determining factor for the success and speed of innovation.

The long-term vision is shaped by three main factors: technology, application domains and policies. Clearly, all factors are drivers of innovation, because (i) anticipated technological advances lead to innovative applications of these advances and (ii) user needs lead to technological innovations that enable applications and services. At the same time, policies and politically established goals and processes lead technologies and applications towards common goals and targets.

Regarding policies, which lead many technologies and applications on a pan-European scale, the ECS community has specified its common objectives that influence and shape long-term innovation and must be considered in future research directions. As presented in the Introduction, these four high-level common objectives are:

- Boosting industrial competitiveness through interdisciplinary technology innovations.
- Ensuring EU digital autonomy through secure, safe and reliable ECS supporting key European application domains.
- Establishing and strengthening sustainable and resilient ECS value chains that support the Green Deal.
- Unleashing the full potential of intelligent and autonomous ECS-based systems for the European digital era.

These objectives, which are aligned with policies and European political priorities, address the need to establish unrestricted access to goods and services, free exchange of know-how and information, under trusted, protected and regulated multilateral agreements in the emerging international political and economic landscape. European Union’s policies to protect its strategic autonomy, and sustain its competitiveness are shaping and continuously advancing, especially for the ECS industry, which constitutes the backbone of the digital society.

European digital strategic autonomy – European Union's ability to maintain control and security of its products, overcoming disruptions and vulnerabilities – is one of the major challenges when considering that its major economic drivers, i.e., digitisation and connectivity, are strongly dependent on the supply of hardware and software from countries outside Europe. This challenge needs to be addressed immediately, for the short term as well as for the long term, by research programs on the following topics:

- Safety and security: development of rigorous methodologies, supported by evidence, that a system is secure and safe; safety and security are requirements for trustworthiness. These methodologies should enable certification through appropriate methods, such as testing and/or formal methods to prove trustworthiness guarantees.
- Artificial intelligence and machine learning (AI/ML): AI/ML-based techniques will contribute significantly to the development of robust ECS components, systems, and applications, with short development cycles. AI/ML will influence all major technologies in ECS development, from model-based engineering and embedded software to fabrication, and will constitute a major link between quality, reliability, safety, and security.
- Trustworthiness: development of methodologies that integrate traditional ECS technologies with AI/ML, from device level up to applications and human interface. Trustworthiness is key to the acceptance of such emerging systems. Advances in explainable AI models for human/ system interaction, safety, security, risk analysis and management, liability and certification are necessary for the required trustworthiness that will lead to the acceptance of the new generation of innovative products.

The European Green Deal is another policy that combines wide civilian acceptance with high political priority and shapes innovation strongly. As climate change and environmental degradation pose an existential threat to Europe and the world, the European Green Deal is the European strategy to make the economy of the European Union sustainable in the long term[1]. By 2050, a modern resource-efficient and competitive economy must be in place, characterized by:

- Zero net emissions of greenhouse gases.
- Economic growth decoupled form resource use.
- Inclusion (no person and no place are left behind).

The ECS community is instrumental to the realization of the European Green Deal. The many challenges associated with energy management can be tackled only with ECS-based solutions, leading to energy-efficient ECS devices as well.

[1] https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en

The first three high-level common objectives of the ECS community (competitiveness, robustness of ECS products and establishing value chains) can be achieved only by reaching the fourth one as well. The "unleashing" of intelligent and autonomous ECS-based systems requires the interdisciplinary effort and coordination of all stakeholders; academic, institutional, and industrial. In the effort to ensure effective and timely identification of effective exploitation of opportunities, a close cooperation of all stakeholders along the value chain is a prerequisite. This cooperation is traditionally strong in Europe and constitutes a valuable European strategic asset. This strength is based on the availability of many research facilities with excellent competence and extensive experience in the ECS domain. This comprehensive ecosystem of universities, RTOs, and industrial research organizations distributed across many countries in Europe forms a leading incubator for pioneering technologies that enable the creation of hyper-smart, safe, secure, and resource-efficient electronic components and systems. This ecosystem enables increasingly networked scientific work and is the base for maintaining the competitiveness of the European ECS industry now and in the future. Cooperation also offers the best opportunities for coping with the growing interdependencies and interdisciplinarity through strong coupling of basic and applied research within the European Research Framework Programme. This, in turn, creates the fertile soil, from which industry can receive substantial impulses to achieve breakthrough solutions with minimal time to market, leading to maintenance of European technological excellence and leadership, which is the cornerstone of long-term European technological leadership and a basis for prosperity and peace in our continent.

Additionally, and independently of policies, long-term vision is shaped by technological and application evolution and revolution. Many future applications will be enabled by enhanced functional and non-functional properties provided by new technologies (both hardware and software), as projected in technology-application roadmaps such as the one shown in Figure 4.1. Typically, the advances that are foreseen through roadmaps are considered evolutionary. However, there have been several occurrences of revolutionary or disruptive developments in technology. These are not projected in roadmaps; they exploit and establish innovative technological models and have tremendous technological and societal impact. Often, they lead to paradigm shifts with significant impact to business and society. The World Wide Web is a typical example of disruptive technology.

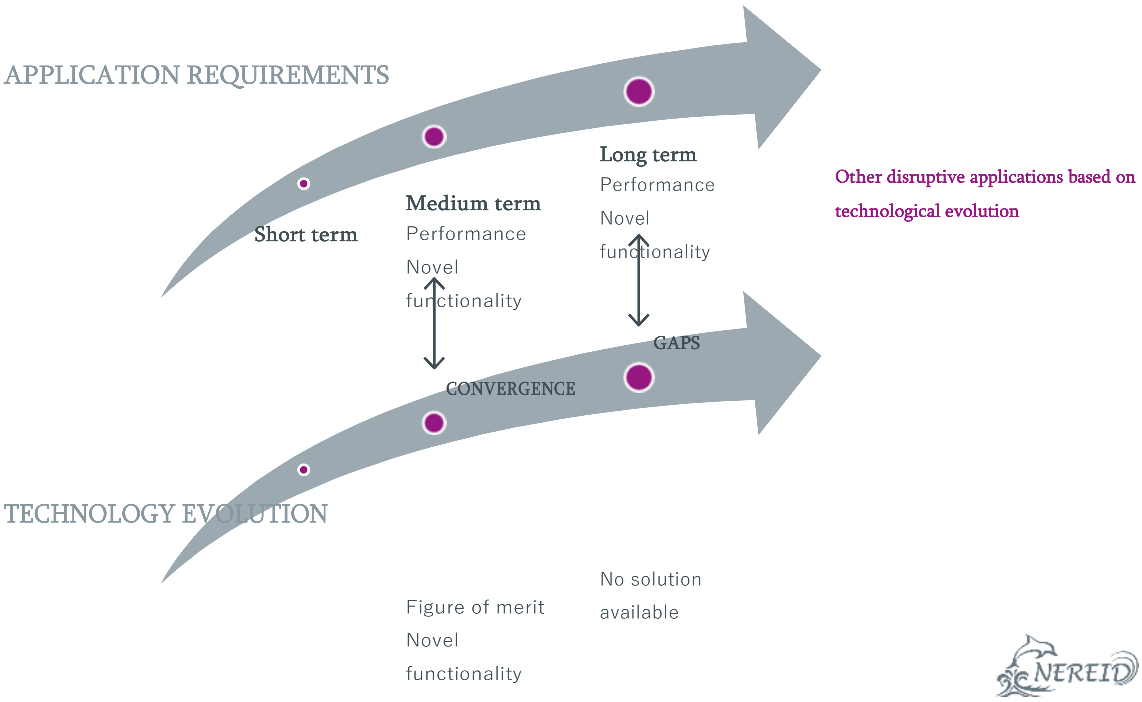


Figure 4.1 - Technology evolution and application requirements

Over the last decades, the ECS domain has evolved from a technology-driven field to an environment where societal needs and application requirements guide the research agendas of the centres of expertise. However, technology-driven research goals need to be a part of the research agendas, considering that novel technologies often create and enable new classes of applications. The European competences in "Beyond CMOS", 'More Moore' and 'More than Moore' have been instrumental in bringing about this change, resulting in a strong European position in markets that require complex multifunctional smart systems. Clearly, maintaining and extending these competences is fundamental to the continuous offering of disruptive technologies that will preserve the European competitive position.

In this Chapter, we present the main research trends that are of particular importance to the European strategic research and innovation agenda. Clearly, presenting a complete list of anticipated evolutionary and revolutionary, or disruptive, technologies and challenges is infeasible, by its very nature. Considering the three factors that shape the long-term vision - technology, application domains and policy - in the following section we present a model that enables us to present challenges in a systematic way. We consider policies to provide the framework as well as parameters for technologies and applications and then, we present technological challenges and needs to be met in application domains.

4.2
MODEL

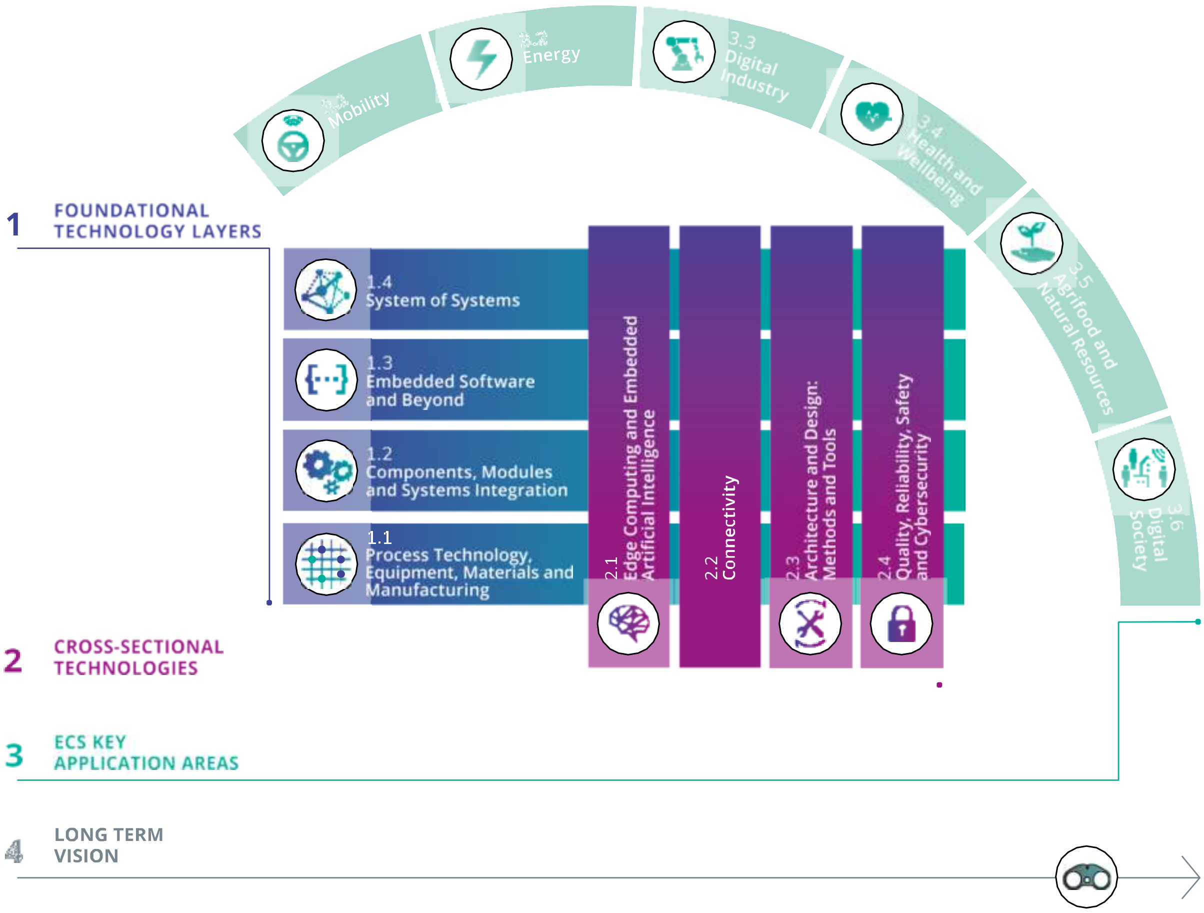


Figure 4.2 - Technology domains evolution and application requirements

As explained in the introductory chapter of this SRIA (Chapter 0), we consider a layered model for the technological and application challenges for ECS, as shown in Figure 4..

In the remaining sections of this Chapter, we present challenges in technologies, fundamental and cross-sectional, as well as in the application domains that are enumerated in Figure 4., with the understanding that our presentation addresses evolutionary and revolutionary technologies based on conventional technological and societal understanding. Independently, our expectation is that disruptive innovations will be readily integrated in our long-term view, since they will affect part or parts of the layered model which we exploit for abstraction, understanding, presentation and openness.

4.3
TECHNOLOGY LONG-TERM CHALLENGES

4.3.1 Process technology, equipment, materials, and manufacturing

Europe has a strong competence in ECS process technology, enabled by the presence of an industrial, institutional, and academic ecosystem with a long tradition in multidisciplinary collaborative research in regional, transnational and European cooperative projects. With the growing complexity of ECS-based devices and systems, this multidisciplinary collaborative approach along the value chain is one of the major assets for Europe in maintaining its competitiveness.

In the More Moore field, there are strong interests in Europe for specific activities that involve very low power devices, leading to possible disruptive applications – for instance, for future IoT systems, devices for AI/ML, neuromorphic and photonic computing devices, embedded memories, 3D sequential integration or application-driven performance (e.g. high temperature operations in the automotive industry).

New materials, including 1D and 2D structures, ultimate processing technologies and novel nanodevice structures for logic and memories are mandatory for different applications as well as new circuit architectures, design techniques and embedded software. Some of these nanostructures are also very interesting for advanced sensors, energy harvesters and photonics. All of these are key for future high performance/ultra- low power tera-scale integration and autonomous nanosystems.

These promising technologies that could underpin numerous future applications will allow us to overcome a range of challenges being faced for future ICs – in particular, high performance, low/very low static and dynamic power consumption, device scaling, low variability, and affordable cost. Many long-term challenges must be addressed to ensure successful application of these nanotechnologies. A number of these are described briefly in the following:

- Nanowires and nanosheets, for high performance and very low-power nanoscale devices, the best material and geometry options for logic (high speed as well as low power) need to be identified.
- Millimetre-wave and THz front ends with III-V MOSFETs have to be developed (with applications in communications, radar, etc.), including 3D aspects of processing.

- Non-conventional switching devices, like negative capacitance field-effect transistors (NCFETs), tunnel effect transistors (TFET), 1D (CNT) or 2D (graphene and others), which could be suitable for very low power devices, need development in basic material, extended characterization of optimal architectures and design strategies.
- For nano-electro-mechanical FETs (NEMS-FET), low-voltage reliable devices have to be developed.
- Spin-based devices also for switching and sensing.

In the field of alternative memories, resistive RAM, magnetic RAM (SOT and VCMA) and ferroelectric RAM/FeFET/FTJ will be key for driving the limits of integration and performance beyond that afforded by existing non-volatile, DRAM and SRAM memories. Research should address:

- Widening the material screening and programming schemes.
- Variability and reliability, especially data retention.
- Trade-off between programming speed and programming power/data retention.
- Compatibility with standard logic processes.
- Architectures for memory embedding in logic, for novel computing schemes.

In the long-term beyond CMOS domain, the challenges to be addressed are in the field of beyond-conventional CMOS technologies, non-Boolean logic, and beyond-von Neumann architectures, including novel state variables, new materials and device, and innovative device-architecture interaction.

Integrated photonics, evolving from silicon photonics, are required to interface conventional electronics with photonic-based communications and sensors, and, in a longer perspective, with photonics-based computing.

The emerging field of Quantum computing poses its own challenges in process technology, equipment, and materials:

- As there are still several candidates for becoming the standard quantum computing technology (such as semiconductor quantum dots, superconductor junctions, photonic circuits, ion-traps, cold atoms, topological states, etc.), a wide range of materials is relevant, together with innovations in process technology.
- New metrology capabilities are required, especially the measurement of electrical properties, such as local carrier mobility, is needed.
- To achieve practical applications, reliable fabrication, connection, and read-out of qubits need to be developed. The low temperatures at which most quantum systems are operated requires the development of cryogenic devices, to interface conventional electronics.

Importantly, all quantum technologies related to sensing, communications and computing, including software, present significant challenges today.

Reducing the environmental impact of semiconductor manufacturing

While the chip advancements are contributing to the industries across verticals, and are crucial to achieve the Green Deal objectives, significant efforts must be made to tackle the direct environmental impact of chips and more generally ECS manufacturing. Issues to be addressed include waste generation, resource usage, CO_2 and GHG emissions, hazardous materials use, including PFAS, and scarce materials use.

Waste generation

Fabricating a small 2g microchip (≈ 14-10nm technology node) requires 32-35 kilograms of water, 1.6kg of petroleum and 72g of chemicals. Since advanced technology nodes will require more metal layers and lithography steps, and, as the chip production could nearly double to satisfy chip demand in the coming years, the environmental impact of the semiconductor industry on power/energy and water consumption, and on CO_2 and GHG emission, will strongly increase. This consumption and emission can reach unacceptable levels to cope with the Green Deal sustainability objectives.

Natural resource consumption

Regarding water, each chip needs to be rinsed with ultrapure water (UPW) to remove various debris (ions, particles, silica, etc.) from the manufacturing process and prevent the chips from becoming contaminated. The semiconductor industry has been working for more than twenty years to reduce the amount of water needed to manufacture a chip for economic reasons as well. Nowadays, due to the most frequent occurrence of droughts, the water issue is a high priority in the sustainable development plans of major semiconductor companies. Semiconductor manufacturers must focus their efforts on new ways to recycle, reduce, and reuse the water used in their production. Nevertheless, new advancements in water treatment must emerge to allow semiconductor manufacturers to recover and reuse wastewater, remove targeted contaminants, and even reclaim valuable products from waste streams. New long-term approaches will deserve more R&D efforts for improving the effluent segregation systems and hence increasing the use of recycled water in semiconductor manufacturing lines.

The main European semiconductor manufacturing should use 100% of renewable energy sources in 2030. Likewise, the water and carbon footprints of the semiconductor industry must be strongly reduced to achieve near-zero CO_2 and GHG emission, for instance.

CO_2 and Green House Gas (GHG) emission

In the semiconductor industry, CO_2 and Green House Gas (GHG) emission arise from process gases used during wafer etching, chamber cleaning, and other tasks. Furthermore, they rise as node size shrinks. These gases, which include PFCs, HFCs, NF_3, and N_20, have high global-warming potential. Gas recycling of unutilized process gases and by-products through various means, such as membrane separation, cryogenic recovery, adsorption, and desorption can be a long-term approach to reducing GHG emission. In collaboration with equipment suppliers, semiconductor fabs could refine them into pure process gases that can be used again, potentially reducing process-gas emissions. For this lever to become economically viable, collaboration between semiconductor companies, equipment suppliers and researchers will be compulsory to address these major challenges related to the separation of process-gas outflows and purification.

Another long-term approach could consist of lowering GHG emissions by switching chemicals that have a lower environmental impact than the aforementioned fluoride gases, such as on-site generation of molecular F_2 for replacing NF_3, since molecular F_2 has no global warming potential. Developing new solutions will require strong R&D efforts and will be both costly and time-consuming, as is the process for qualifying new chemicals on existing processes and tools.

Since most of the aforementioned fluoride compounds are used for etching, another long-term approach could concern the replacement of some non-critical etching processes by additive manufacturing process steps. Such a replacement will require strong R&D efforts to develop highly selective deposition processes and/or self-assembled molecules that can prohibit the deposition of metal and dielectrics.

Sustainability issues in semiconductor manufacturing induced by PFAs

PFAS is a class of thousands of synthetic substances known as 'forever chemicals' since they do not break down in the environment. Most of which are either persistent themselves or are transformed into persistent compounds in the environment. These substances are hazardous for human health as they accumulate in the body, but also in water, ground, and then the seas and oceans.

PFAS are defined as fluorinated substances that contain at least one fully fluorinated methyl or methylene carbon atom (without any H/Cl/Br/I atom attached to it), i.e. with a few noted exceptions, any chemical with at least a perfluorinated methyl group (–CF3) or a perfluorinated methylene group (–CF2–). Due to the C-F bond strength compared to C-C bond, PFAS offer a unique set of technical characteristics, which include exceptional heat and chemical resistance, high electrical insulation resistance, high purity, low-outgassing and low coefficient of friction.

Those intrinsic properties are the basis of many of the technical benefits of fluorinated materials in semiconductor processing, but this also leads to their chemical stability and environmental persistence. Fluorination brings unique physicochemical properties and consequent qualitative improvements that are the enablers of semiconductor, performance and manufacturing advancements.

PFAS are used in the semiconductor manufacturing industry in the lithography process, as a component added to the photoresist to generate photo-acid generators (PAG), improve its adhesion to the silicon wafer, increase its durability, and enhance its resistance to harsh chemicals and high temperatures. In addition to their use in photolithography, PFAS are crucial in producing other semiconductor components. They are used in wet chemistries as surfactants (cleaning, stripping and etching, and metal plating), dry etching, chamber plasma cleaning, CVD and ALD. PFAS are paramount in packaging materials to improve thermal stability and moisture resistance. They are also used as a coolant in the chip etching process, as working fluids for vacuum pumps…

Besides their application in chipmaking, PFAS are also essential for semiconductor manufacturing equipment and factory infrastructure. Their exceptional properties, such as heat resistance and chemical inertness, make them useful in equipment components (tubing, gaskets, containers, filters, etc.) and lubrication (such as various oils and greases).

Up to now, excluding some dedicated applications, there is no PFAS-free alternative for most of the aforementioned applications.

As of today, 1485 tons of PFAS are used every year for producing semiconductors in the European Economic Area. Since the European Chips Act aims at doubling the EU's current manufacturing capacity from 9-10% to 20% by 2030, it will require at least a four-fold expansion of the semiconductor manufacturing capacity in the EU, and consequently a four-fold use of PFAS.
Accordingly, it will be crucial to conduct research to, where possible, identify alternative chemistries that are preferable from an environmental point of view and to develop measurement, recycling, treatment, and efficient abatement technologies to prevent environmental releases for uses to which no PFAS alternative can be found.

Scarce materials use

Finally, to secure their whole supply chain and for not wasting mineral resources, in view of the limited extractable quantities of metals in the earth's crust, chipmakers will be increasingly concerned with the potential scarcity of some ores that are compulsory for producing ultra-pure metals for the high-volume manufacturing of devices. One approach to resolve this issue is to use recycled metals instead of premium metals. Another one is to recover the metals from the electronic waste (e-waste), thus, preventing the use of natural resources. The recovery of scarce metals from microelectronic devices opens a wide research domain for material scientists, ensuring sustainable metal sources for chipmakers.

4.3.2 Components, modules, and systems integration

The interaction among people and other information agents and their environment usually features a trade of data, which is curated into information that either results in a gain of knowledge and/or the enabling of purposeful action or reaction to a given situation.

The width and breadth of such data trading is expected to increase in the future in terms of space and time density. Seamless integration / interaction with the environment and agents involved, based on evolved human-machine interaction (e.g. haptic and brain-computer interfaces) or machine-to-machine interaction, is expected in scenarios that either empower or substitute humans in decision loops.

ECS conform to the HW and SW ensembles that at different levels of integration and organization complexity, mediate the different elements of such information trading: acquisition, management, and exploitation. Indeed, HW and SW integration schemes are the ones ultimately responsible for substantiating the expected increasing number of systems functions and applications that are to emerge from the reunion of:

- New sensitive and structural materials.
- Physical-to-digital (and vice versa) transducers architectures.
- Local (on-system) intelligence.
- Communication/interaction interfaces with users or higher instances of the decision chain.

In addition to the continuous improvement of semiconductor processes and materials, increasing the level of proficiency of elements managing information, integration of more diverse components will be essential to make systems aimed at monitoring the condition of people, assets, processes, and environments less dependable on the use of energy and on external supervision. Making these devices and systems faster, more sensitive, efficient, robust, functional, and apt to different application scenarios will demand higher levels of heterogeneity of materials and fabrication and assembling processes.

Self-powering, energy harvesting and storage will become more and more important and significant advances are expected in solid-state devices to cover the needs of edge and IoT devices. Integration of intelligence to these inherently power-restricted devices requires novel power-efficient computational platforms, such as neural networks and analog computing approaches in parallel with CMOS and other traditional semiconductor devices as commented in the previous chapter.

Next generation computing devices, using physics to make computation, pose challenges in integration as well as in development. In such approaches, envisaged in Chapter 3.1, and made possible in the frame of processes of Chapter 2.1, other modes of coding information besides bits will be used, e.g. using qubits or encoding in time, like for neuromorphic architectures where information is coded in a succession of spikes, or their coincidence in time. Another massively parallel approach using biological technology (based on proteins, DNA construction, etc.) can also emerge for niche applications, or for storage[2]. Most of these technologies will be used first in servers for very specialized acceleration but will slowly improve to be integrated into edge devices.

Moreover, ECS are pivotal elements of the digital transition that supports the current and future quest for making our civilization sustainable, maximizing performance and minimizing e-waste, and particularly for slowing down, reverting or making human environments resilient to climate change.

In addition to the sustainability of the ECS fabrication processes themselves, the scarcity of materials and the increasing demand for ECS imply that approaches and architectures that improve the chances of ECS modules to be repaired and/or reused, and of their material constituents to be reclaimed, need to be addressed. Single-use or disposable devices need to be designed with minimal, and minimally invasive, electronics. The same applies to install-and-forget systems to be deployed in natural remote locations. For example, organic and printed electronics can lead to biocompatible electronics and more effectively recyclable systems. Although there are widespread efforts to recycle computing systems, we are far from the goal of effectively recycling because of the lack of processes to support an appropriate circular economic and business model. Component and system level challenges range from homogenizing component and subsystem lifecycles to computing models and materials used. In a wide range of application domains, a significant challenge is the ability for global reconfiguration of system resources to satisfy diverse applications' functional and non-functional requirements, such as latency and energy, including re-training in AI/ML subsystems. Significant effort needs to be made to develop systems that are scalable linearly or functionally.

In terms of global resilience to climate change, multifunctional smart information systems will be in demand to react faster to such upcoming challenges and risks. For instance, globalization of human activity and large-scale weather changing patterns will ease the spread of known diseases beyond their usual geographical boundaries as well as spur the appearance of new ones. Swinging weather conditions will affect application fields directly exposed to climate conditions such as mobility, energy, or agrifood/ environmental applications. ECS helping these applications to react to those abrupt changes need to be integrated and packaged themselves in a way that can cope with these harsher environmental conditions.

Integration, as the art of recursively combining physical devices, components, and systems together to form a new entity with increased functionality in the minimum volume possible, will be key to leverage the different positive aspects of diverse technologies and their reference materials. It has been already appraised that a combination of nanoelectronics, photonics (optoelectronics), electronic smart systems, including AI/ML subsystems, and flexible, organic and printed electronics is setting the path for future enabling functional electronics[3], which will be characterized by aspects such as:

- A shift from physical to functional integration.
- The use of novel substrates and structural systems.
- Seamless integration in everyday objects for a broad spectrum of new applications.

[2] https://www.microsoft.com/en-us/research/project/dna-storage/
[3] https://5e-project.eu

- Real-time capture and management of multi-physics data and contextual information.
- Safe and secure operation.
- Networked, autonomous operations complemented by software solutions (including AI).
- Eco-design approaches at product, process, and business model levels.

The distinction of monolithic and heterogeneous integration, and what can be achieved with them, is subjected to boundaries that will evolve with time. Particularly, monolithic integration at chip, chiplet, and SoC levels is progressing through the development and maximum exploitation of 3D sequential integration, a technology with important research activities in the EU that will impact applications with very high-density interconnections (IoT, neuromorphic computing, etc.). Heterogeneous integration, from non-CMOS materials and device processing on top of CMOS wafers to customized application platforms, is also progressing thanks to the evolution of scalable wafer-level or package-level integration schemes nurturing compact System-in-Package (SiP). Still, maximum versatility comes with integration of technologically dissimilar components (e.g. MEMS/MOEMS- NEMS and ICs, electronic and photonic elements, etc.) onto application-oriented platforms, which could be board-like or built on flexible/conformal substrates. To serve such versatility, 3D place-and-route tools with extended ranges of speed, precision, and gentleness for handling components that on occasions are fragile, will be needed. Modelling/simulation, characterization and reliability evaluation tools, which are also strong European domains, will be required to take all the new materials, technologies, device architectures and operation conditions into account, so that cost of development is reduced, and technology optimization is speeded up. All those encompassing schemes are expected to be beneficial for the integration of future high-performance sustainable, secure, ubiquitous, and pervasive systems, which will be of great added value for many applications in the field of detection and communication of health problems, environmental quality, secure transport, building and industrial monitoring, entertainment, education, etc.

4.3.3 Embedded software and beyond

The next generation cyber physical systems will play a key role in the future AI, IoT, SoS realisations, while they will need to be sustainable and easy to maintain, update and upgrade in a cost-effective way, across their complete lifecycle. Mature software platforms running on them will ensure safety and security by design and be available as a part of the European digital infrastructure to a wide audience for building services and business.

We envision an open marketplace for software frameworks, middleware and digital twins with a seamless integration and ubiquitous presence that will represent a backbone for the future development of one-of-a-kind products. While such artefacts need to exploit the existing software stacks and hardware, they also need to support correct and high-quality software by design.

Thus, the envisioned long-term achievements in embedded software will drive the digital industry, while enabling collaborative product-service engineering, and making sure to be inclusive.

For overcoming challenges related to efficient engineering of software, new programming languages and tools for developing large-scale applications for embedded SoS will emerge. Software engineering will address hybrid distributed computing platforms, including efficient software portability, and the development of new software architectures involving edge computing will follow. Model-based testing will contribute to handle uncontrolled SoS.

Short-delivery cycles, maintenance, and extension of software systems are goals that require continuous integration and deployment. Autonomous embedded systems and autonomous processes for IoT & edge embedded HW/SW co-design, and integration & orchestration platforms for IoT and SoS will contribute to achieving those goals. Model-based engineering, based on multi-dimensional, complex and of scale digital twins in the edge, as well as their deployment along with systems, will contribute to a continuous integration. Next generation hybrid digital twins, based on big data-driven and classical physics principles, will be developed and integrated in embedded hardware, while supporting enhanced cognition and intelligence that will demonstrate enhanced capabilities to encapsulate the real word, e.g. power modules, while enabling unseen capabilities for supporting high-level missions as the Green Deal.

As anticipated, software in cyber physical systems must support sustainability: approaches for lifecycle management will enable this by supporting distinction between core systems capabilities and applications and services, and by enabling interplay with legacy subsystems. Interoperability must be built-in and ensured by integration platforms, and will enable features such as easy SW updates, device management, and data management. Composability of systems will be a property supported by properties contracts and orchestration systems, and will be directed towards “write once, run anywhere” for optimal execution on the cloud-for-edge computing continuum.

The modularity of future cyber physical systems, which will dynamically compose in large SoS, will require a high-level of trust, both at the level of the constituent components of SoS and considering how they connect and compose. To this regard, the evolution towards stronger protocols and interfaces (including well-defined pre-compilation connections), supporting security, privacy, and dependability aspects, represents a key factor.

Use of safe, trustworthy & explainable AI will be dominant in autonomous systems and will enable embedded intelligence. AI will play several key unconventional roles in innovation, e.g. as a tool for SW development/ engineering. These innovations will be supported by the European Processor Initiative and its integration in cloud servers, open-source hardware, and software.

Finally, use of quantum computing and IoT digital twin simulation will support software reliability and trust.

4.3.4 System of Systems

System of Systems (SoS) is projected to become an area of exceptional economic growth, both short term and over the coming decades[4]. This will create a strong market pull for the complete ECS value network upstream of the SoS area.

Strategic investments addressing open platforms, engineering, and deployment efficiency, SoS management and control represent key factors to propel Europe towards very large scale of digitalization and automation solutions across integrated and optimised operations of engineering, production, logistics, infrastructures, etc. The inherent heterogeneity of SoS is expanded with the new, emerging computational models that include accelerators, AI/ML subsystems, approximate computing, organic systems, and others, pose significant challenges at all levels. Interoperability and adaptation to diverse physical interfaces and communicated data structures constitute clear examples of challenges. The management of heterogeneity at all levels, including dynamic instantiation of multi-paradigm computing resources considering application requirements and specifications, auto-configuration of distributed resources (locally or globally) to satisfy application functional and non-functional requirements, require significant advances in the field. Success in this direction requires additional activities towards standardization for HW/SW functions as well as scalability specifications to achieve wide and cost-effective use in applications and usecase with variable performance requirements.

Large scale usage of SoS technology is further expected to be a significant contributor to the Green Deal through distributed and intelligent solutions that provide significant reduction of environmental footprint in terms of energy consumption, material consumption, waste and, in general, through a more rational and controlled use of all types of resources. This strengthens the ECS value network through energy efficient and robust electronics hardware, connectivity, and embedded software.

[4] Advancy, 2019: Embedded Intelligence: Trends and Challenges, A study by Advancy, commissioned by ARTEMIS Industry Association. March 2019. Available online at: https://artemis-ia.eu/publication/download/advancy-report.pdf.

The future evolution of SoS will further require cooperation between domains, enabling a wider shared understanding of the context and situation, more useful services, richer functionality, better user experience and value proposition. This evolution will introduce the concept of connected and interacting domains (potentially both physical and virtual), where application and services run transversally on top of connected vertical domains.

In the medium/long term, many technologies will allow the evolution of SoS towards the scenarios previously described, including:

- Distributed AI, to control the inherent and quickly increasing complexity of SoS, making them secure, reliable, easier to maintain, etc.
- Connected and interacting domains, supported by:
    - Open and robust integration platforms.
    - AI method adopted to address conflicting functional and non-functional requirements.
- Engineering support for emerging behaviours in complex SoS:
    - Model based engineering.
    - Predictability, controllability, monitoring and diagnoses.
- Automated and autonomous engineering.
- Machine interpretable content.

4.3.5 Edge Computing and embedded artificial intelligence

Artificial intelligence will be the enabling foundation for the digital society, ensuring that the systems that make up its framework function in an effective, efficient, secure, and safe manner. Most of the ambitions that are to be realized in the digital society, such as a zero-emission economy, affordable healthcare for everyone, safe and secure transactions, etc., can be achieved only if an underlying AI infrastructure is in place. This implies that the Internet of Things will gradually transform into the Artificial Intelligence of Things (AIoT), where AI constitutes the interface between the digital world (e.g. edge and cloud computing, cognitive and autonomous cyber-physical systems, embedded systems) and the analogue real world.

Artificial intelligence and machine learning (AI/ML) methods enable efficient and effective automated decision making in domains ranging from system design, design space exploration and manufacturing to application and business processes. As computing models distribute functionality at all systems from the cloud to the edge, AI/ML methods need to be distributed and coordinated, leading to efficient smart systems at all levels of the computing hierarchy. In addition to the ongoing research in AI/ML for applications in increasing application domains, efficient and effective methods for distributed intelligence and federated learning become increasingly important. Advanced AI approaches, like composite AI, require heterogeneous technologies to be addressed altogether, such as vision and natural language processing. The accelerating adoption of AI/ML, in its various approaches, results in high demand of subsystems and accelerators, creating a significant set of challenges analogous only to the growth of the respective market. Importantly, considering the known social questions regarding the adoption of AI, significant effort needs to be spent on certifiable and explainable AI, which will lead to the necessary social acceptance of AI-related technologies at all fronts.

AI technology is becoming increasingly demanding for computational power, especially for the learning phase. As Figure 4. shows, the need for increased accuracy in AI techniques leads to methods that employ increasing numbers of parameters (for deep learning techniques, specifically, in Figure 4.), which, in turn, lead to dramatic increase of need for computational power to implement these techniques. This will imply new progress in energy efficiency to keep the Cost of Ownership affordable. Mainly the GAFAM and BAITX will be able to afford the computing infrastructure that will require a large number of servers.

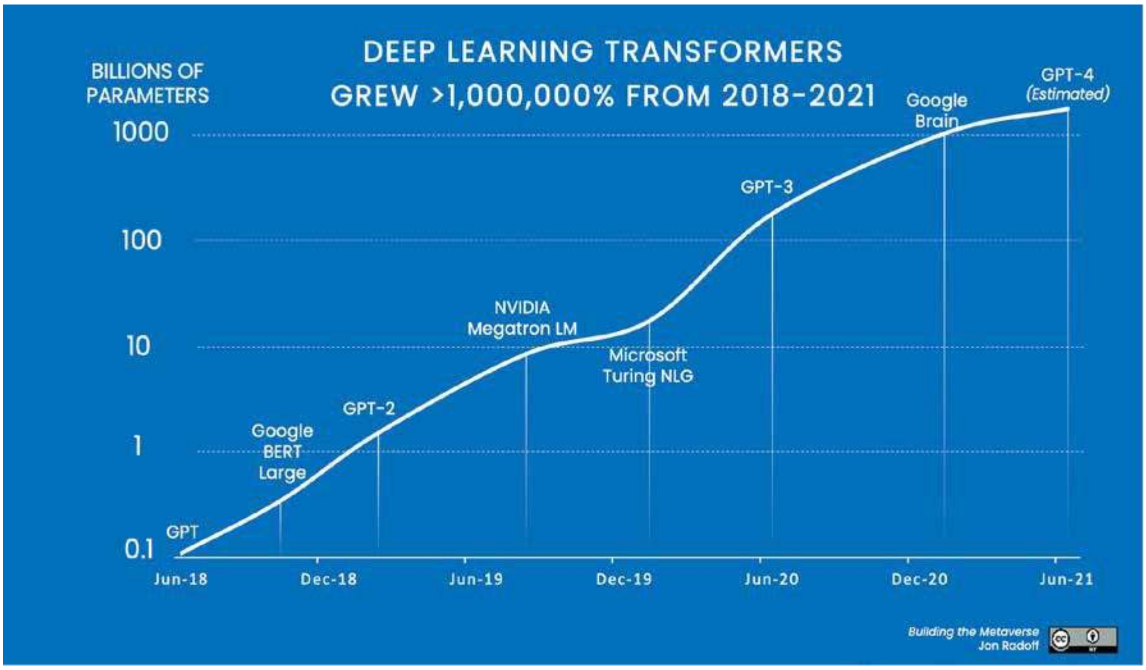


Figure 4.3- The increase of parameters employed for deep learning systems (2018-2021)

IA accelerators will appear in many devices from the deep edge to home servers, allowing to process all kind of data and changing the way we interact with computers. Figure 4. demonstrates this, showing the dramatic increase of deep learning chiplets that have been shipped worldwide in the last years. Computing systems will disappear in the environments and will allow natural interactions.

To be able to cope with the diversity of requirements and the Cambrian explosion of designs, AI techniques will be used to select the best architecture (automatic design space exploration) and to generate the code from high-level specifications (no-code) with guarantees of correctness. AI/ML application in the ECS domain include the automated design of SoCs, the respective design space exploration, the integration and orchestration of multiple computing paradigms into embedded systems (including AI-based ones) as well as the design and evaluation of cyber-physical systems overall, which include physical components for operation or computation.

New AI paradigms will emerge, including self-supervised approaches, that can be highly efficient at the edge and decreasing the need for a large database and computing power like for deep learning. Quantum computing, in addition to other computational models, needs to be explored for efficiency and effectiveness in AI methods.

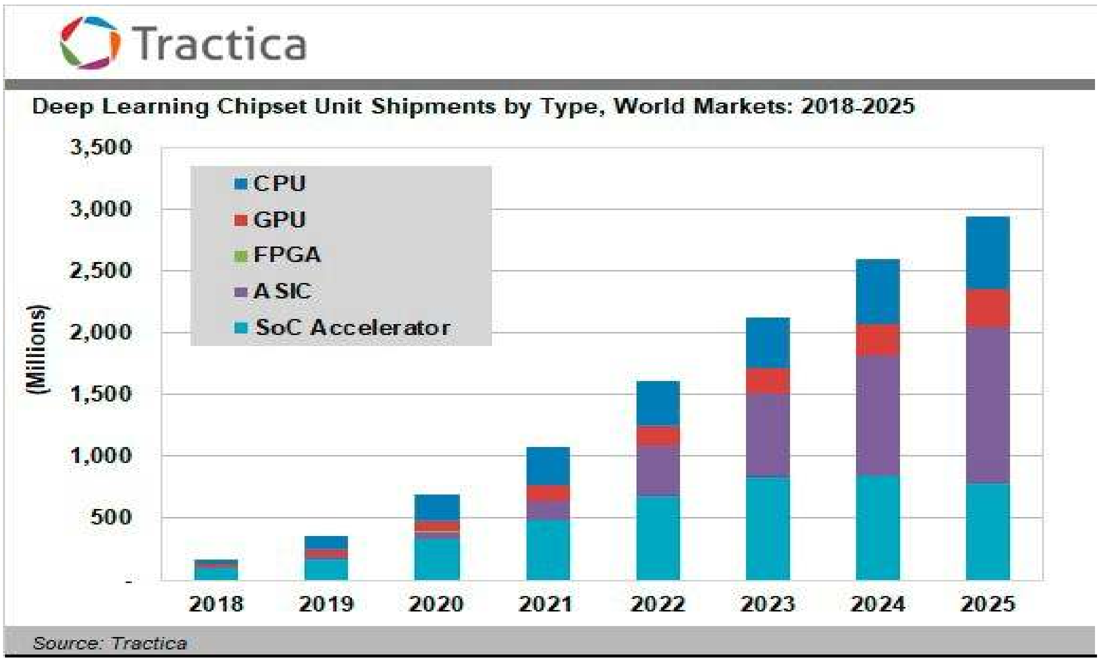


Figure 4.4- The growth of deep learning chip shipments (2018-2025)

The mass market for smart systems will still be fulfilled by very advanced foundries, but "small" foundries will also appear (assuming that they become sufficiently cheap and efficient) to make the very diversified deep edge devices that most of the time don't need the latest technology. Some low-cost applications will be made with non-silicon technologies, e.g. using printing or 3D printing technologies[5]. This will allow enterprise level organizations to have their own "device manufacturing" that don't need high quality clean rooms. AI/ML technologies will play an important role in this direction, towards high automation that leverages AI/ML for the exploration of architectures and code generation.

4.3.6 Connectivity

Today, connectivity is a key enabler to support the development of innovative applications in several markets, such as consumer, automotive, digital manufacturing, network infrastructure (to name a few). This is also pushed by the need of being connected anywhere+anyhow+anytime. The availability of innovative connectivity technologies, both wireless (e.g. low-power wide area networks (LPWAN),

[5] https://www.nature.com/articles/s41586-021-03625-w

cellular (5/6G)) and wired (e.g. new bus-oriented communication protocols), will enable and enhance a wide range of new business opportunities for the European industry in the context of Systems of Systems (SoS), Cyber-Physical Systems (CPS), and Internet of Things (IoT). Long-term roadmaps for connectivity and interoperability will guide a seamless integration of heterogeneous technologies (hardware and software) for the design and implementation of complex connected systems in effective ways.

Connectivity is a critical asset to any digitalization and automation activity to strengthen Europe's position and enable the European industry to capture new business opportunities associated with the connected world we live in. It is vital to support European technological leadership in connectivity, fostering digitization based on Internet of Things (IoT) and System of Systems (SoS) technologies; for example, this can be achieved by being at the forefront of new standard development for the current 5G initiative, the emerging SoS market, and the upcoming 6G initiative. Furthermore, to bring added value and differentiation with respect to US and Asian competitors, The European industry has to secure access to any innovative software and hardware technology that enables the efficient engineering of large and complex SoS (which will help to capture more value by targeting higher-end or more innovative applications, as highlighted by the Advancy report[6]). For instance, connectivity (e.g. in terms of wireless infrastructure market, led by Ericsson and Nokia) will be supported by the European leadership position in the traditional IT environments as well as the embedded segments, guaranteed by companies such as STMicroelectronics, Infineon and NXP. Connectivity from device, over the edge, and to the cloud will need to be virtualized, relying on run-time design, deployment, and management of integrated edge and cloud network architectures. This will enable the connectivity from cloud to far edge, shifting the perspective from point-to-point connectivity to application-to-application connectivity.

Connectivity engineering and management must be significantly improved to support simplified and inexpensive deployment and integration of new applications into SoS, CPS and IoT solutions.

Connectivity will provide the basis for a data layer supporting instant and seamless data and information exchange between producers (supply below the data layer) and consumers (demand above the data layer) within and between domains. This layer will enable large-scale integration of SoS, CPS and IoT solutions. Targeting systems and applications, we should consider the interconnection between sub-systems and focus on individual component technology development, according to needs identified at system or application level. To support this system vision, the promotion of innovative technology enabling heterogeneous integration is key. To fully leverage this heterogeneous integration at hardware level, software interoperability is a parallel challenge to provide connectivity that will allow for SoS integration. Thus, an alternative major challenge is to enable SoS integration through nearly lossless interoperability across protocols, encodings, and semantics. To do so, dedicated software tools, reference architecture and standardization are key to supporting SoS integration, thus enabling the provision of a scalable and evolvable SoS. As it remains very difficult to assume that highly customized embedded systems will be built based on a single, unified, high-level modelling principle and toolset, there is a quest for consolidation, or even the standardization of basic

---

[6] Advancy, 2019: Embedded Intelligence: Trends and Challenges, A study by Advancy, commissioned by ARTEMIS Industry Association. March 2019. Available online at: https://artemis-ia.eu/publication/download/advancy-report.pdf.

runtime frameworks, component libraries and subsystem interfaces that will ease the deployment of interoperable components into generic, domain-specific solutions and architectural frameworks in a bottom- up fashion. Such an approach is also expected to provide for better traceability of requirement validation, and formal verification of distributed system compositions and their emerging functional and non-functional properties.

Finally, data protection must be ensured at an appropriate level for each user and each functionality, regardless of the technology. One major challenge is to ensure security interoperability across any connectivity. This foresees the utilization of different connectivity technologies, and these differences create security incompatibilities leading to increased engineering costs. Therefore, the development of innovative hardware and software security solutions, that will support and provide correctness and safety, is of fundamental importance. Such a solution will have to be linked with the previous challenges to ease SoS engineering, deployment, and operation in a seamless manner. Security assessment is a significant issue here considering the criticality of applications. Standards and directives are required not only for technology transfer and system evaluation, but for legal purposes as well, considering the existing GDPR legal framework and the emerging laws regarding European and national cybersecurity requirements.

Thus, the following major challenges need to be addressed in the connectivity roadmap until 2050:

- Keeping European leadership in connectivity.
- Providing virtualized connectivity.
- Introducing data-oriented connectivity.
- Developing connectivity engineering.
- Meeting future connectivity requirements leveraging heterogeneous technologies.
- Enabling nearly lossless interoperability across protocols, encodings, and semantics.
- Ensuring secure connectivity and interoperability.

4.3.7 Architecture and design: methods and tools

The European ECS industry has been strong in systems engineering, integration, validation and verification, test and simulation, and certification of innovative ECS-based products. The produced systems are characterized by high quality in such terms as functionality, safety, security, reliability, trustworthiness, and certifiability.

Considering the need to maintain and increase this strength, we need to invest in extending existing, and developing new processes, methods and tools that will ensure European leadership in the field. Emerging ECS components and systems are characterized by new functionalities, increased complexity, and diversity on all fronts, ranging from methods and paradigms to modelling and analysis. As the competition by US and Asian ECS companies is fierce, significant effort and investment is necessary to enable European leadership in the technologies for integration, validation, verification, testing and certifiability. A significant parameter in the establishment of leadership and effective technologies is the support of the European Green Deal, by enabling green development and green ECS-based products. Tools and methods for managing the complete ECS lifecycle are necessary, ranging from resource-considerate and climate-neutral design and operation to development, production and maintenance of ECS-based products addressing issues that include even decommissioning and recycling.

To realize this vision and associated goals, the European efforts need to extend existing, and develop new, processes and methods that cover the whole lifecycle of products, from initial requirements elicitation through design, integration, verification, validation, test, certification, production to commissioning, operation, maintenance, and decommissioning. These processes and methods need to support data collection from production as well as from operation and maintenance to be analysed and used for continued development and integration, updates in the field, validation, verification and test at the development phase, as well as in the field, at run-time.

Novel architectures and development and analysis tools need to be developed which will enable:

- Seamless design, development, integration, verification, validation and test across all layers of the technology stack, from semiconductor up to systems of systems. These methods need to address individual ECS-based systems, groups of systems that form and dissolve statically or dynamically as well as systems that cooperate with other systems and with humans, at the cloud or at the edge. Furthermore, methods and tools need to support open platforms and integration of open systems.
- Verification, validation, and test of highly automated and autonomous systems, especially coping with open-world assumption and uncertainty.
- System development that includes AI methods, such as explainable and trusted AI.
- The use of AI-based methods in the design and development process, for example design space exploration and analysis, including certified products.
- Managing the increasing functionality, connectivity, and complexity of systems.
- Managing the increased diversity of tools, such as modelling and description languages and simulation and testing tools, in emerging components, modules, and systems.

4.3.8 Quality, reliability, safety, and cybersecurity

Quality, reliability, safety, and cybersecurity are fundamental components of any innovation in the digital economy. Especially in Europe these characteristics are particularly important since European products are well-known to be of high quality in almost every aspect. They are driven by high expectations of the European society demanding these features. Continuous evolution of our European society is driven by the development of electronic components and systems (ECS). Advances in mobile systems and applications enable one with a mobile phone to can buy a flight ticket, make a money transfer, or maintain social contacts; new services are becoming continuously available. ECS simply promises to make our lives more comfortable, safe, and efficient but this promise relies on user trust and acceptance concerning the perception of sufficient privacy, security, understandability, and usefulness in daily life. In the near future, highly automated and autonomous systems supported by AI will have a constant growing trend. We expect that in the next ten years such systems will be increasingly deployed, not only in controlled environments, such as in manufacturing industries, but massively spread in our personal, professional, and social spheres.

ECS of the future will not require an external environment control to work as wished. More generally, the ECS of the future will have to satisfy different constraints on different scientific and social disciplines and ought to meet both the founding principles of European society.

To maintain European leadership in electronic devices and systems we must make our efforts to provide innovative products of the well-known high quality, reliability, safety and cybersecurity to our customers. From this perspective we expected the following challenges to be considered in a long term:

- Development and integration of new materials for advanced packaging and interfaces, new characterization techniques, and new failure modes caused by new use-case scenarios.
- AI/ML methods, including digital twins, to be a cornerstone of keep leadership regarding quality and reliability of ECS made in Europe and be an enabler for new data-driven business models.
- Model-based engineering (incl. standardization of data management and processing) to be a key instrument for virtual release of ECS through the supply chain and shortening time to market.
- Assuring user trust and acceptance of ECS through early inclusion of user requirements, explainability-by-design, and user education and training.
- Model-based engineering (including standardization of data management and processing) to be a key instrument for virtual pre-qualification of ECS and shortening time-to-market.
- Data transmission methods and protocols that are so reliable that they can be used to transmit life-sustaining information over long distances (e.g. for robotic surgery).
- Reliable and certified software that can be kept even if the underlying hardware or hardware architecture is changed, including potential influence of SW updates on HW

reliability.

- Software that can adapt to a degrading underlying hardware to achieve a long-lasting and reliable HW/SW combination.
- Liability of trusted AI-driven systems (it is based on trustworthiness of (AI-driven) systems, included safety and certification of AI-driven CPS, which is a main challenge in 2021).
- Safely manage /design for human interactions in complex systems, SoS and application scenarios
- Ensuring sustainability, cybersecurity, safety, and privacy, for AI-driven and quantum-based systems (based on "ensuring safety, security and privacy and sustainability of (AI-driven) systems").
- The attention to the environment and privacy have increased significantly in the population. This implies that conceiving any safety solutions is not enough. Our vision on 2030 is an integration between disciplines, which have nothing to do with safety or computer science in stricto sensu, such as privacy, social trust, liability, and sustainability.

Importantly, all properties of quality, reliability, trustworthiness and safety are heavily dependent on the integrity of the supply chain that is employed to develop ECS and ECS-based systems. The business models that are employed within Europe and abroad, lead to dependence of organizations on suppliers and contractors for provisioning of hardware as well as software components and systems. Trust among them is not a given. Many cybersecurity incidents that have hit headlines originate from exploited vulnerabilities in the supply chain. Conventional verification and testing methods are not sufficient to address these problems, which range from insertion of hardware trojans at fabrication plants and implanting malicious hardware components in systems to compromised system and application software. Addressing these problems requires significant research effort for new methods to validate systems at all levels of hardware and software.

4.3.9 Machine Learning and Artificial Intelligence

The recent advances in machine learning (ML) and artificial intelligence (AI) have opened several opportunities in hardware design and design tools, especially in embedded computing and design of cyber-physical systems. The major directions are: (i) exploitation of AI/ML techniques in component and system design, (ii) architectures and designs for efficient AI/ML processing in emerging solutions, and (iii) effective AI/ML methods for embedded and cyber-physical application domains. Thus, AI/ML is increasingly becoming a fundamental technology that influences all foundational technology layers and cross-sectional technologies depicted in the model shown in Figure 4.2.

The increasing complexity of integrated circuits (IC), still following Moore's Law, leads to a (still) exponentially increasing design space that needs to be explored for IC designs at all levels of design abstraction. The problem of design space exploration is not new and has been addressed with various methods in the past, in the context of design synthesis. The problem includes identification of designs that meet design specifications as well as optimization of desired parameters, such as delay, area, power, etc. Heuristic algorithms have been employed extensively to identify appropriate designs, since exhaustive search of design spaces for modern IC's is infeasible. AI/ML methods are increasingly adopted in design space exploration in place or complementing heuristic algorithms at several stages of circuit design, from manufacturing and physical design to RTL design and high-level synthesis. Importantly, AI/ML are increasingly exploited in design testing and verification.

AI/ML methods are used in several approaches to address different problems in the system design process. A comprehensive review of the problems and approaches is given by M. Rapp et al.[7], which classifies the problem types solved with ML with 3 parameters: (i) development of predictions, actions and/or data, (ii) the design stage and (iii) the exploited AI/ML algorithm. In terms of predictions, currently, AI/ML methods are used and proposed for predicting design properties after explored decisions, while for actions, they are used in place of existing techniques at various stages of design, such as RTL design, placement and routing, etc., with the purpose to optimize an appropriate criterion. As most of the employed algorithms are supervised learning ones, data constitute a significant component of effective employment of AI/ML techniques. As such data do not exist or may differ significantly in practice, because of the different existing tool chains, AI/ML methods are used to produce or collect training data within and from the design process itself. Employed algorithms range from regression to neural networks, mostly for supervised learning, based on models that range from graphs to images at the various stages of the design process.

Despite the large and increasing research work in employing AI/ML in embedded system design, there are several challenges that need to be addressed for the penetration of research results in practical design frameworks. The biggest problem is that of appropriate data for training the models; this is a well-known problem in most application domains of AI/ML. The need for large amounts of appropriate data, considering the exponential space of design alternatives that need to be evaluated, will lead to long delays until the appropriate models are built. Alternative methods for generating appropriate date will also need to be developed.

Machine Learning (ML) and, more generally, Artificial Intelligence (AI) algorithms are essential to extract relevant information from massive data amounts. As the computational capabilities of edge- and end-devices (also indicated as “far edge”) is ever increasing, ML-based inference processing is no longer carried out in the cloud, but is distributed across the device-edge-cloud continuum. This requires the identification of a proper architecture to efficiently distribute the computational load, with dynamic methods for model compression or model splitting among the architecture tiers. Edge AI or, more generally, edge intelligence has emerged as crucial in multiple applications. Likewise, the possibility of implementing tiny ML (TinyML) algorithms in constrained devices opens new perspective in the Internet of Things (IoT). Distributed ML will involve several aspects, including latency, bandwidth utilization, data safety, data quantization, privacy, cost.

AI/ML accelerators constitute innovative components in new generation computing systems of all scales, from large and core to edge. Research in this area is continuously growing with accelerators

[7] M. Rapp, H. Amrouch, Y. Lin, B. Yu, D.Z. Pan, M. Wolf and J. Henkel, “MLCAD: A Survey of Research in Machine Learning for CAD.” In IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems, 41(10), Oct. 2022.

that target application domains, from low-power embedded systems for inference to back-end high-performance accelerators for model training. Several surveys exist presenting methodologies for AI/ML accelerations, especially for neural networks[8,9,10] , while there also exist surveys of commercial accelerators[11]. The wide range of application domains of accelerators poses significant challenges to designs, from performance to power to cost. Especially for embedded systems, accelerators are effectively targeted for edge systems, from automotive to medical, from industry to consumer. From the surveys it becomes alarmingly clear that most technologies are developed and commercialized by organisations outside Europe. European efforts exist, such as AnIA[12] , Axelera AI's AI Edge accelerator, Infineon's Parallel Processing Unit (PPU) AI accelerator[13] and others. However, market penetration, even in the start-up domain, is limited when compared to the US and Asia. Efforts focus on specific application domains, such as computer vision, automotive, etc. Strengthening of these efforts is necessary, to address challenges for cost-effective accelerators in a wide range of domains. The challenges include development of effective architectures and exploration of the trade-offs among memory design, power consumption, high performance and arithmetic calculation precision that dominate AI/ML neural network accelerators, in conjunction with specialized algorithms.

Significant challenges also exist in adopting AI/ML in embedded and cyber-physical application domains. AI/ML methods are already used and are continuously being developed for effective operation of embedded and cyber-physical systems in various environments, through predicting operational parameters, including input, and/or environmental parameters. However, often, the unpredictability of parameters combined with the lack of effective training data leads to suboptimal operation and creates a significant challenge to develop effective solutions that are application dependent.

4.3.10 Quantum Technologies

Quantum technologies are promising to provide effective solutions to a wide range of application domains. Some application-specific systems are at the first steps of commercialization, such as quantum sensors for gravitometers and neuroimaging and quantum communication devices specifically for quantum key distribution. At the other end, quantum computing requires significant research and advanced development effort to produce large scale, fault tolerant and general-purpose systems. Considering the challenges and the characteristics of quantum systems, the expectation is that quantum computing will complement conventional computing and digital systems, focusing on specific complex problems such as fractioning large numbers, optimizations

---

[8] V. Sze, Y. Chen, T. Yang, and J. S. Emer, "Efficient Processing of Deep Neural Networks: A Tutorial and Survey." Proceedings of the IEEE, 105(12), Dec. 2017, pp. 2295–2329.

[9] E. Wang et al., "Deep Neural Network Approximation for Custom Hardware." ACM Computing Surveys, 52(2), May 2019, pp. 1–39.

[10] F. P. Sunny, E. Taheri, M. Nikdast, and S. Pasricha, "A Survey on Silicon Photonics for Deep Learning." ACM Journal on Emerging Technologies in Computing Systems, 17(4), Oct 2021.

[11] A. Reuther, P. Michaleas, M. Jones, V. Gadepally, S. Samsi and J. Kepner, "AI and ML Accelerator Survey and Trends." 2022 IEEE High Performance Extreme Computing Conference (HPEC), MA, USA, 2022, pp. 1-10.

[12] IMEC, "AnIA: a novel AI IC." Available online: https://www.imec-int.com/en/expertise/cmos-advanced/compute/accelerators#ania

[13] Synopsys, "Infineon and Synopsys Collaborate to Accelerate Artificial Intelligence in Automotive Applications." Available online: https://news.synopsys.com/2019-09-17-Infineon-and-Synopsys-Collaborate-to-Accelerate-Artificial-Intelligence-in-Automotive-Applications

requiring multiple variables, solving quantum problems as in quantum chemistry and physics, data analysis and sorting, and data encryption.

Quantum systems have been developed exploiting the wave-particle duality at the physics level and their ability to achieve discrete states. Their power originates from the exploitation of their superposition and entanglement properties. Superposition refers to their ability to exist in two or more states at once, while entanglement refers to their ability to create interdependencies among particles even when separated by large distances.

Superposition and entanglement enable powerful operations in terms of computing and communications, considering that entanglement promises high speed and secure communications while superposition coupled with entanglement enables effective parallel processing. Quantum networks and a Quantum Internet are the targets of several research efforts, including a test site in Chicago in the USA. Quantum computers are the focus of several efforts around the world, from large companies, such as IBM, Google and Intel, and small companies, like Rigetti, IonQ and Oxford Quantum Circuits, to universities such as TU Delft in Europe and the University of Science and Technology of China.

Quantum operations in computing and communication systems are based on quantum bits, named qubits, where a qubit constitutes the smallest unit in which quantum information is generated, transduced, processed, stored, and transmitted. A qubit is a superposition of the classical binary states (0,1). Analogously to classical computers, quantum computers perform a series of operations (a quantum algorithm) to modify qubit superpositions (probability of being in a particular state) and entanglements to increase some probabilities and to reduce others. Measurement of a qubit causes its state, and the states of entangled qubits, to collapse to either '0' or '1' with a probability dependent on the state of the qubit at the time of measurement. The goal is to maximize the probability of measuring the correct answer.

There are two main quantum computing models today:

- Analog or adiabatic quantum computing, and
- Gate-based quantum computation.

Analog quantum computing is typically based on quantum annealing, which performs processing by initialization of the system followed by slow, global control of the qubits towards a final state and readout. In this computational model, the energy state of a quantum system encodes or models a problem and the energy landscape, which starts flat, changes slowly and continuously to a final state with the energy peaks and valleys representing the problem to be solved. Analog computers constitute some of the most developed quantum systems to date, but their functionality is limited to simple and very specific problems, because they have limited ability to reduce noise, which impairs qubit quality.

Gate-based quantum computation is analogous to traditional gate-based computers, using a sequence of quantum gates, each composed of a few qubits, that perform logical operations, followed by measurement. Unlike many classical logic gates, quantum logic gates are reversible. A universal set of quantum gates is still needed to achieve the full capability of quantum computation. Gate-based quantum computing is error prone, due to noise. High gate and qubit error rates limit the scalability of quantum systems, leading to the challenging requirement for fault-tolerant quantum computation (FTQC) that would enable system scalability. In this direction, gate error rates as well as qubit and gate fidelity are metrics for the robustness of gate operations; qubit fidelity measures the loss of qubit coherence due to its interaction with its environment and due to shifts of its quantum states over time. These metrics essentially measure how closely actual gate operations match -on average- ideal versions of these operations. Analogously to classical computing error correction methods, quantum error correction (QEC) dramatically lowers effective error rates by encoding the quantum state using redundant "physical" qubits and using a QEC code to emulate stable qubits with very low error rates, often called "fault-tolerant" or "logical" qubits. Importantly, logical qubits currently require a large number of physical qubits and many quantum gates to maintain their state, incurring significant resource overhead in terms of both additional qubits for each "logical" qubit, and additional quantum gates for each logical operation. As an example, for simulating a chemical structure, a FTQC computer based on 111 "logical" qubits would require 10^8 and 10^5 "physical" qubits when the physical qubit error rate decreases from 10^{-3} to 10^{-9} [14]. The development of an accurate and general-purpose quantum computer would need between 10^4-10^6 "logical" qubits (Fig. 1) and a quantum and gate fidelity around ($1-10^{-15}$) for industrial applications[15]. Since QEC leads to considerable overhead with conventional technology, Noisy Intermediate Scale Quantum (NISQ) computers are an intermediate step towards large scale error corrected fault-tolerant quantum computers (FTQC). NISQ computers use qubits without QEC, and although not fully fault-tolerant, they are expected to be practical in the near term for some applications. NISQ computers will require around 10^2 to 10^4 logical qubits (Fig. 1) and a very high qubit fidelity, improving the mean error rate of current quantum computers which reaches 4%. To reduce the impact of environmental noise, and of the unwanted interactions on qubit fidelity, quantum error mitigation (QEM) algorithms are used to mitigate these errors rather than completely remove them.

Quantum computers are expected to find their first use cases in combination with classical computers, so that the QC will serve as an accelerator, that solves the one part of the computation that is computationally expensive for a classical computer, and the classical computer runs other tasks. In this way, even a relatively small and noisy quantum computer can provide benefit in a larger problem.

Figure 4. shows the expected evolution of the number of "logical" qubits for both FTQC and NISQ approaches.

---

[14] National Academies of Sciences, Engineering, and Medicine. 2019. *Quantum Computing: Progress and Prospects*. Washington, DC: The National AcademiesPress.

[15] B. de Jong, How about quantum computing?, Berkeley labs. Available at: https://cs.lbl.gov/assets/CSSSP-Slides/20190624-deJong.pdf [Last access: Oct. 15, 2023)

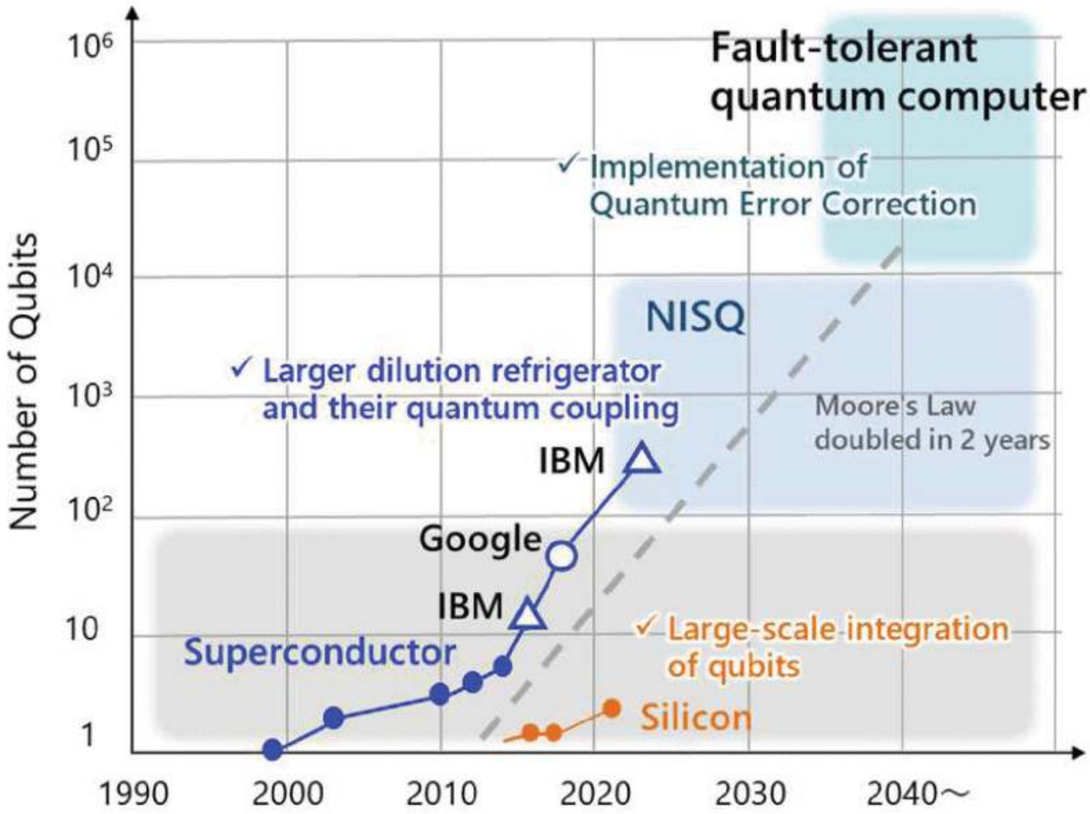


Figure 4.5 - The trend of the number of "logical" qubits and the goal for NISQ and FTQC approaches

The number of qubits, and qubit and quantum gate fidelity in a system are the defining parameter for effective quantum computing. For this reason, significant research and development effort is spent on investigating alternative technologies for qubit manufacturing. Current efforts focus on superconducting qubits, semiconductor gate-defined quantum dots, color centers, trapped ions, cold atoms, photons and topologically protected Majorana modes. These alternative qubit technologies differ significantly in physical aspects, operational and miniaturization challenges. Neutral atoms in vacuum are entirely different from superconducting and electron spin qubits, and photon-based qubits, also named "flying qubits", are also quite different from other types of qubits which are static in location; in contrast to solid-state qubits, photons do not have significant decoherence problems but are harder to generate, control and detect in a deterministic way. Figure 4. provides a mapping of current qubit technologies in terms of maturity and research intensity, while Table 1 presents the properties of the alternative qubit manufacturing technologies.

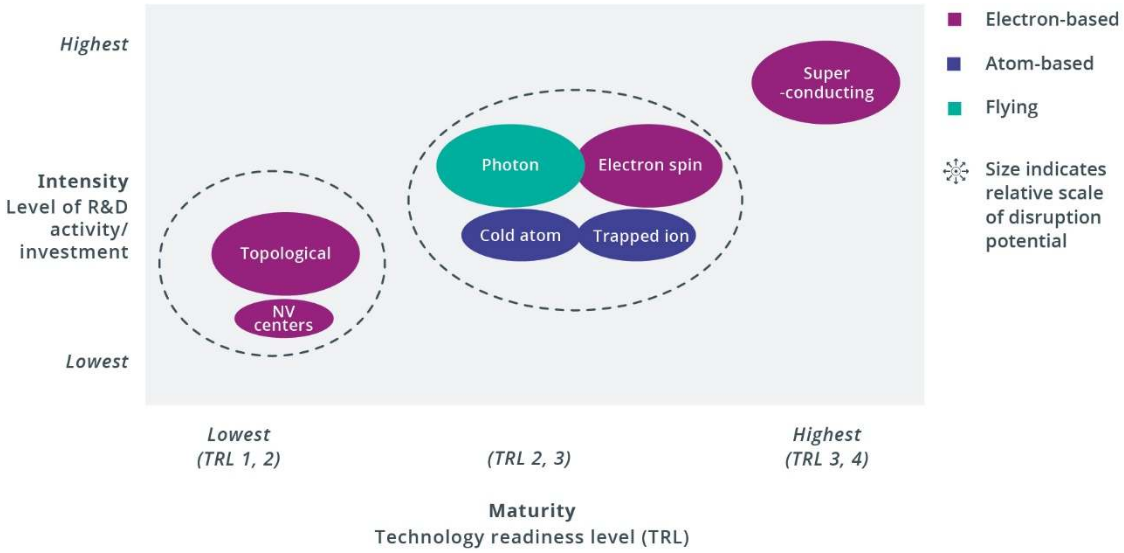


Figure 4.6 - Core qubit technologies mapped by maturity, intensity, and disruption potential

Table 1: Properties of various qubit manufacturing technologies for quantum computing

Qubit technology	Superconductor	Trapped ion	Si e^- spin	Cold atom	Photon	Color centers
Qubit size	100μm²	1mm²	100nm²	1-10μm²	5-25μm²	100nm²
Quantum gate	Microwave Magnetic field	Laser Microwave	Magnetic field	Laser Microwave	Interference	Microwave
1-qubit fidelity	99.96%	99.999%	99.93%	99.9%		99.9952%
2-qubit fidelity	99.3%	99.9%	> 99%	99.5%	98%	99.2%
Gate speed	12-400ns	100μs	1μs	0.4-2μs	1ns	20-50ns
Coherence time T_2	150μs	50s	20ms	40s	150μs	0.6s
Variability	3%	0.01%	0.1-0.5%	-	0.5%	large
Operation Temperature	15mK	10K	1K	4K	4-10K	≈ 273K
Entangled qubits	433	32	6	256	≈ 20	5

The different quantum computing approaches lead to challenges in problem-solving aspects of quantum computers, in addition to the manufacturing issues. Quantum algorithms and software for quantum computing are in their first steps and present significant challenges. The goals of user-friendly quantum computers that are programmed effectively and efficiently require significant advances in terms of algorithms and software tools. Recent advances in specialized application domains such as localization, optimization and machine learning exploit quantum computers effectively but require very specialized knowledge. Popular current environments and tools, such as Qiskit, demonstrate the ability of tools to make quantum computing more attractive to new generations of engineers, but the challenge of producing tools such as appropriate compilers and interpreters, needs to be met to enable a sufficiently wide adoption and evolution of quantum computing.

4.4 APPLICATION EVOLUTION AND LONG-TERM CHALLENGES

4.4.1 Mobility

The European Union has issued ambitious policy statements regarding transport and smart mobility:

- Emissions from transport could be reduced to more than 60% below 1990 levels by 2050.
- The EU has adopted the Vision Zero and Safe System approach, to eliminate deaths and serious injuries on European roads.
- Sustainable Mobility for Europe: safe, connected, and clean.

To realize this vision, possible scenarios include the projection that mainly autonomous and electrically driven vehicles (FEVs) will be on the road, and that all road users will be connected. It is envisaged that other road users (bicycles, pedestrians, public transport) will also participate in this connected, autonomous model, in addition to transportation network infrastructure (tolls, signals, etc.), creating an augmented Internet of Vehicles. Key networking technologies, such as the emerging 5G cellular connections with their very low latency (ms range) and the powerful edge nodes (Mobile Edge Computing, MEC), will enable highly effective vehicular communications for traffic management and safety applications. Railways and maritime transport will also become more autonomous. Fully integrated multimodal traffic will be applied, in which air, railways and maritime are fully integrated with road transport.

Until now, the rule was simple: more income equals more travel distance as *Figure 4.7* indicates. Will this very simple equation still work in the future?

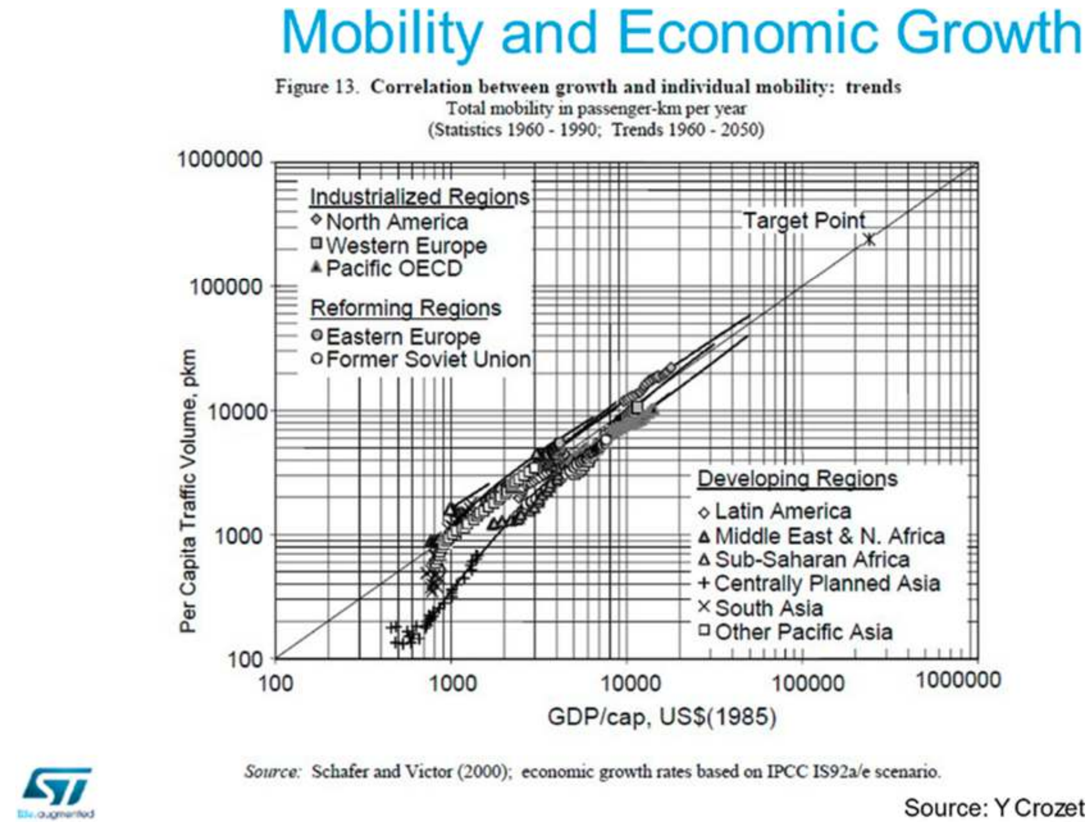


Figure 4.7- The correlation between growth and individual mobility

Many attempts have been made in the past to move people from individual to public transport. Huge investments have been made into infrastructure, fast trains have been deployed massively, bus lines installed. Nevertheless, this did not change the distribution of shares in the transport between planes, trains and cars. It could just allow to keep pace in an ever-increasing mobility. The most important means of transport is still, and will also very probably be in the next decade, the individual car.

If we want to reduce emissions and energy waste, then we need to focus on this means of transport and make it more ecologically friendly. It must use less space and energy.

The global mobility is undergoing a significant change triggered by the increasing thread of global warming. The European Union started the Green Deal, which is radically changing the way mobility works. Classical fossil fuel-based transport and mobility will be completely replaced by CO_2 neutral mobility. As all the alternative mobility systems, be it battery-based, H_2 based or synthetic-fuel-based, have their individual challenges, all of them have in common that CO_2 neutrality is only possible in using a connected, shared and energy-usage minimized mobility network. To reduce the space needed by cars, the best solution would be to use them better during the day by sharing their usage.

Owning cars should thus be replaced by shared mobility and intermodal transportation offering the most convenient and CO_2 neutral way to move goods or persons from point A to point B. This will be only possible by a stable, everywhere available, user-friendly, secure, fast communication system connecting people, traffic operations, cars, trucks, busses, airplanes, ships, busses, etc. across the globe.

Today's vehicles contain more software than any other embedded system and most compute applications. Tomorrow's vehicles will multiply the software lines of code by a factor 6. Semiconductor value in the car will more than double through the next 10 years (from 600$ to 1200$). Some already talk about ''software- enabled vehicles" or "data-centres on wheels". Besides the electronics and software to get the car rolling, there will also be a lot of complexity added in terms of performing safety & security checks and to monitor the health or lifetime of electronic components and batteries. Overall, the evolution of the car and mobility in general results in the rise of complexity in electronics and software that has almost become uncontrollable.

Additionally, the growing global population as well as the aging society will be supported by more and more automated transport means at all mobility variants taking the best advantage of the available resources as roads, parking space, airspace, and water space, and serving best the needs for mobility of the society. This will be supported by sensors combining different sensor principles in one sensor with significantly less power consumption, as well as new AI optimized edge computers in the transport vehicles. These systems will be part of completely new HW/SW systems spanning from sensors via embedded edge computers via predictable, fast, clean (also for the user), safe, secure, and failsafe communication to globally interconnected cloud systems.

4.4.2 Energy

Power electronics is the enabling technology for the efficient generation, conversion, distribution, and usage of electrical energy. It is a cross-functional technology covering very high gigawatt (GW) power (e.g. in energy transmission lines) down to the very low milliwatt (mW) power needed to operate a mobile phone, and even to microwatt (μW) to power autonomous sensor nodes. Many market segments, such as domestic and office appliances, computers and communication, ventilation, air conditioning and lighting, factory automation and drives, traction, automotive and renewable energy, can potentially benefit from the application of power electronics technology. The ambitious goals of the EU to reduce energy consumption and CO_2 emissions can only be achieved through extensive application and use of power electronics, as this is the basic prerequisite for:

- Efficiently feeding wind and solar energy into the power grids.
- The stabilization of the power grids with an increased share of fluctuating renewable energy sources.
- Highly efficient, variable speed, motor drives.
- Energy-efficient and low-emission mobility with hybrid and full electric vehicles.
- Energy-saving lighting technology.
- Efficient recovery of braking energy.
- Energy management of batteries.
- Control appliances and building management systems via the grid interface (smart grids).

The estimated energy savings that can be achieved by introducing state-of-the-art and future power electronics components into systems is enormous, estimated at more than 25% of current electricity consumption in the EU. Since power electronics is a key technology in achieving a sustainable energy society, the demand for power electronics solutions will show significant growth over the coming decades. The European industry holds a strong position in the field of silicon-based power semiconductors and modules and is establishing a robust foundation for future progress in wide bandgap semiconductor technology. Europe also has high-quality power electronics research groups at universities and research institutes with well-established networks and associations across Europe to provide platforms for discussion, cooperation, and joint research.

A long-term roadmap for power technology needs to cover different sectors.

- New, highly efficient power devices based on wide-bandgap semiconductor materials such as SiC and GaN-on-silicon, and possibly Ga_2O_3, AlN, diamond, diamond-on-silicon, or nanowire-based materials.
- New, cost-efficient, Si-based power devices to enable high efficiencies for mass-market applications such as super-junction MOSFETs.
- Power management for very low-power applications as required for IoT, including the development of energy harvesting technologies, covering the full range from GW to μW levels.
- High temperature-capable packages serving new materials and 3D technologies that offer the highest requirements and integration capabilities.

In the energy roadmap towards 2050, five major challenges were identified:

- Smart & Efficient – Managing Energy Generation, Conversion and Storage Systems, trying to fulfil the vision of loss-free energy conversion and generation.
- Energy management from On-Site to Distribution Systems.
- Transmission grids with the goal to achieve in 2020 solutions to cope with rising grid loads as a base for the carbon-free energy transition of Europe.
- Efficient Community and Regional Energy management.
- Cross-Sectional Tasks for Energy System Monitoring & Control, so that highly integrated monitoring and control of energy systems and grids, are achieved utilizing innovative ECS-based solutions.

These challenges need to be addressed to achieve the current EU policy target of 30% savings by 2030 by utilizing innovative ECS-based solutions, as well as the milestones of (a) -55% GHG emissions until 2030 (getting closer to zero emissions due in 2050) and (b) grid integration. To realize this vision, we need to target the decentralization of energy sources, opportunities with networked systems, limitations in peak electricity supply, oversupply times, new demand for electric energy supply for urban mobility, and the introduction of storage systems. This will lead to new challenges in energy management providing flexibility, stability and reliability in the grids and distribution for communities and cities. Furthermore, we need to develop components for HV transmission, of 1.2 MV or even higher voltages, to roll out an efficient energy transmission over Europe. Also, we need to combine local generation & demand site management with transmission & distribution grid operation & control technologies from sub-MW to GW scale, and we need to develop resilient solutions coping with adverse conditions resulting from the advancing climate change.

Additional technical solutions are needed to increasing share of renewable energy generation, self-consumption (mainly heating/cooling and EV) and building optimization, as well as introducing and managing new types of renewable energy carriers like hydrogen.

Relevant promising technologies, already under use and extension, include (a) artificial intelligence & advanced communication techniques for cyber-security increasing resilient energy system control, and (b) optimal control of distributed generation and dispersed energy-storage devices as well as robust, high power control devices.

4.4.3 Digital industry

Digital Industry is a must of European productive and commercial evolution on the next decade, following and empowering the EU policy related to digitalization. Digital capabilities and functions will be the enabler for safer, greener, sustainable, lower cost and more productive, autonomous, and competitive EU industrial ecosystem.

EU planned, mostly after the Covid experience and lesson learnt, strategic investments addressing digitalization including industrial productive arena, edge technological engineering studies, developments and deployment efficiency and services, addressing, in addition to industrial production, logistics, transportation, health, critical EU infrastructure etc.

The future evolution of EU Industry into Digital Industry will further require cooperation between multiple domains enabling a wider shared understanding of the context and situation, filling the gap towards EU industrial strategic autonomy and more useful services, richer functionalities, better user experience and value proposition introducing the concept of connected and interacting domains.

The manufacturing industry can essentially be classified into two main categories: process industry and discrete product manufacturing. The process industry transforms material resources (raw materials, feedstock) during a (typical) (semi)continuous conversion into a new material that has significantly different physical and chemical properties than the starting substance. Discrete manufacturing refers to the production of distinct items. Automobiles, furniture, toys, smartphones, and airplanes are examples of discrete manufacturing products. The resulting products are easily identifiable and differ greatly from process manufacturing where the products are undifferentiated, for example oil, natural gas, and salt. Another meaningful way to distinguish between manufacturing industries is by dissecting the domain by the end-product categories, such as energy industry, chemical industry, petrochemical (oil & gas), food industry, pharmaceutical industry, pulp & paper industry, steel industry (process industries), and furthermore car manufacturing, machine industry, robotics, and the semiconductor industry. Also, these subdomains constitute significant industrial domains for Europe. These industries are ever more demanding and voluminous consumers of ECS technologies such as sensors, big data, artificial intelligence, real-time system, digital twins, safety & security, computing systems, lifecycle engineering, human-system integration etc. ECS technologies are essential parts of most of the advances in these domains.

The perspective of industry is reflected in several efforts. The major ones are described in the following:

- The SPIRE[16] Roadmap 2030 and the SPIRE Vision 2050, which lists the following targets.
  - Replacement of fossil-based materials by bio-based materials requiring completely new processes.
  - Re-use of waste streams that require complete redesign of materials, products, and related production processes.
  - New resource efficient applications that require completely new designed processes.

[16] https://www.spire2030.eu/what/walking-the-spire-roadmap/spire-Roadmap

- Complete redesign of industrial parks to realize industrial symbiosis.

- The Factories of the Future (EFFRA[17]) roadmap summarizes its vision as follows:
  - Agile value networks: lot-size one – distributed manufacturing.
  - Excellence in manufacturing: advanced manufacturing processes and services for zero- defect and innovative processes and products.
  - The human factor: developing human competences in synergy with technological progress.
  - Sustainable value networks: manufacturing driving the circular economy.
  - Interoperable digital manufacturing platforms: supporting an eco-system of manufacturing services.
- The Connected Factories[18] project forecasts the emergence of new manufacturing concepts, such as:
  - Hyperconnected factories.
  - Autonomous factories.
  - Collaborative product-service factories.

Recently, a federation was set up of the three electronics ecosystems in Europe in nanoelectronics, electronic smart systems and flexible, organic & printed electronics (https://5e-project.eu/). Combinations of nanoelectronics and flexible organic & printed electronics that provide functionalities to electronic smart systems lead to novel solutions. This trend to functional electronics is characterized by the following aspects:

- A shift from physical to functional integration.
- The use of novel substrates and structural systems.
- Eco-design approaches at product, process, and business model levels.
- Real time capture & management of multi-physics data and contextual information.
- Networked, autonomous operations complemented by software solutions (incl. AI).
- Seamless integration in everyday objects in a broad spectrum of new applications.

All these efforts provide high-level targets, which translate into diverse and much more concrete targets in each domain, ending up also in a number of technology challenges in this ECS-SRIA, such as distributed AI along the edge-to-cloud continuum, computation and simulation capabilities, communications and interacting domains, engineering support for emerging behaviours in complex SoS, model-based engineering, predictability, controllability, monitoring and diagnosis, automation, autonomy and robotics, teleoperation, telepresence, simulation and training.

[17] https://www.effra.eu/factories-future-roadmap
[18] https://www.effra.eu/connectedfactories

Clearly, ECS technologies that enable distributed Industrial IoT (IIoT) systems to monitor and control manufacturing systems and processes will enable disruptive industrial innovations and realise the vision of Industry 4.0 and the Industrial Internet that will lead manufacturing worldwide. Overall, these long-term trends translate into the need to invest in technology research and innovation projects in the following areas:

- The rise of artificial intelligence (AI), a powerful edge, and cloud computing networks; methods and algorithms need to evolve to more complex, reliable and explainable AI.
- Collection of measurement data, including image, video, and 3D animation, and, in general, large volumes of heterogeneous and unstructured data.
- New production schemes such as:
  - Modular factories, i.e., smaller standard units to be assembled according to needs, also mobile units.
  - More end-user -driven agile production, i.e. end-users more connected to production and logistics chains.
  - Hyper-connected factories.
- New production technologies, e.g. 3D printing, and other novel emerging methods, leading to production that is closer to customers.
- Methods to extend closed-loop production lines to closed-loop regions (extensive recycling, net energy, zero-emission and waste, close to end-users).
- Autonomous to human-machine co-work, to enable flexibility and reduce excessive complexity.

Recyclable electronics, since the digital industry will increasingly become a producer and enabler of “green electronics over the next decade”, leading to the need to recycle as many electronic components and systems as possible.

4.4.4 Health and wellbeing

The rising cost of healthcare, caused by an aging population, is one of the major challenges that present-day society must deal with. To keep healthcare accessible and affordable for everyone, it will change radically in the coming decades. Healthcare will become increasingly decentralized and personalized, as medical care will move from the hospital to people's homes as much as possible. This transition in healthcare can only be achieved through the massive development of digital healthcare devices that can provide personalized monitoring, mentoring and treatment.

ECS will keep on being key enablers to realize the continuum of healthcare, notably in linking well-being, diagnostics, therapeutic approaches and rehabilitation issues. In addition to providing the tools for personal management of individual health and monitoring of health condition, ECS and smart systems will play an active role in assistive technologies with the goal to reduce inequalities linked to impairments originating in loss of physiological or anatomical structure or function after a disease or an accident. Ambient Assisted Living (AAL) is a high-priority direction for Europe, to support its increasing aging population.

In the long term, personalized and patient-tailored healthcare will be at the forefront of technology advancement. Further miniaturization of biomedical devices and integration of smart integrated systems (e.g. smart catheters, electroceuticals) will have significant impact on point of care diagnosis and treatment. Real-time localized detection of disease and minimally invasive targeted drug delivery will be a key priority. Achieving enhanced reliability and building stakeholder confidence in these technology advancements will be key to successful implementation. Data integrity and security around the use and storage of personal information will require new methods of application development and a robust system of operation, especially if moving towards a more connected healthcare approach with more focus on tailored patient diagnosis and treatment.

Beyond those technological challenges, aspects such as reliability, safety and privacy issues in terms of regulation and uptake by practitioners, especially when dealing with procurement policies, must be tackled. A priority will be in bringing these stakeholders closer in the involvement phase of developing key enabling technologies (KETs) for healthcare applications with a customer pull and technology push approach.

Improvements in medicine over the ages greatly benefited from advancements in other disciplines. Medicine evolved over time from a "mechanical" medicine (surgery) toward "chemistry" medicine and more recently biotech medicine. Nowadays, the development in ICT and digitization has an important impact on the way healthcare is addressed. In ten years from now "digital medicine" will be deployed, and will complement, not necessarily replace, the tools offered to medicine to improve the benefits for patients and medical professionals.

These tools may include, for instance, human models also known as the “digital twin”. Here, ECS will have a crucial role in ensuring the necessary link between the digital and the real twins. Real time acquisition and processing of data and vital parameters collected from on-body IoT sensors, is a key technology that will advance existing wearables and will enable identification and prediction of a person’s condition. The use of AI technologies, based on extended measurement data, will enable significant advances in this area.

Finally, progress in interfacing electronics components and systems with biological systems will offer seamless connection to the body for continuous monitoring but also for electrostimulation purposes. Results from the human brain flagship project will provide input for improved deep brain stimulation. Electroceuticals and nerve stimulation will enhance treatments of diseases and partially replace pharmaceutic treatments, thus avoiding side effects.

Some additional developments are presented in the following:

- Fully personalized medicine will be enabled by smart monitoring of health parameters, including factors from the molecular to the environmental levels. Developments in healthcare will benefit from the concept of “digital twins”, so that prediction of health evolution and preventive treatment will become reality and standard procedures. Fully personalized and accurate health data will be available anywhere, anytime.
- Drug development will be assisted by emerging methodologies such as ‘organ-on-chip’.
- 3D-bioprinting. Medicine is highly benefiting from advancement in other disciplines such as genomics or 3D printing. Combining 3D printing of living material and of electronic systems will develop a bottom-up approach to medicine, with advanced and personalized prosthetics and implants increasing biocompatibility, solving the problem of powering, and increasing quality of life.
- Cyborgisation. Future Brain-Computer Interface (BCI) technology will enable new ways of communication, e.g. for people with severe disabilities. By the 2040s wearable or implantable BCI technology will probably make smartphones obsolete. Due to the massive exposition of the physical and biological world in cyberspace, BCI systems will have to incorporate new means of protection of technology, data, and consciousness – like heartbeat, venous system, fMRI or 'Brainprints' as the top measures of security.

These innovations in the medical domain can be accelerated by the creation of an ECS-based technology platform for medical applications. The Health.E Lighthouse[19] initiative has compiled a list of emerging medical domains where further technical developments are required:

- Bioelectronic medicines.
- Organ-on-Chip.
- Personal ultrasound.
- X-ray free interventions.
- Smart minimally invasive instruments.
- Smart drug delivery.
- Intelligent wound care.
- Ambulatory monitoring.
- Point-of-care diagnostics.
- Remote sensing and monitoring.
- E-health.

Despite this urgent need and the enormous resources that are being invested in research, true innovation in terms of products reaching the market has been slow. One of the root causes identified is the lack of open technology platforms. This will release the power of Moore's Law, that has been the driving force in electronics for more than fifty years, to the healthcare domain: "Moore for Medical". It is the vision of the Health.E Lighthouse that innovation can be accelerated by stimulating the development of truly open technology platforms.

The list of challenges that ECS will face in the next decade is changing and new issues, linked to the developments described above, will have to be addressed. Security and reliability remain major issues to guarantee safety and integrity of medicine. Regulation will have to be developed to address these concerns. Furthermore, ethical issues may become more and more critical in the uptake of patients and may lead to fundamental decisions in the way medicinewill evolve.

[19] https://www.health-lighthouse.eu/emerging-medical-domains

4.4.5 Agrifood and natural resources

Over the following decades the global population will increase, rising to an estimated peak of 9.78 billion by 2064. By the middle of the century, about two-thirds of the population will live in urban areas. This will require new digital approaches to supply the growing number of people with food, which will involve a great threat to food security for certain countries and especially for large cities. Digitalization has already helped initiate open field farming through precision agriculture, but there are other ways of targeting this issue, especially by the emerging areas of “digital farming” and “vertical farming”. In this form of farming, plants are grown in vertical arrays, inside buildings, where growing conditions can be optimized. Crops are supplied with nutrients via a monitored system under artificial lighting and can thus be grown year-round. This method makes it possible to grow plants without soil and natural sunlight, with optimal growth conditions being created artificially. The full potential of this approach can only be achieved with the help of information technology (IT) and IoT components and paradigms such as AI and Industry 4.0, which all still need to be adopted for this purpose. With these digital farming approaches, it will be possible to secure food supply autonomy and food safety for large parts of the EU. Furthermore, investigation into the provision of corresponding technologies and approaches will enhance the strategic autonomy of Europe.

The European Green Deal is a response to these challenges, considering the huge negative impact by the climate change. It includes two main programs “From farm to fork” and “Biodiversity 2030” having a strong impact in the goals of this Chapter, which should contribute to reach the targets defined by these two programs by the introduction of the adequate ECS technologies and solutions.

From Farm to Fork

European food is already a global standard for food that is safe, plentiful, nutritious and of high quality. Now European food should also become the global standard for sustainability. EU agriculture, the manufacturing, processing, retailing, packaging, and transportation of food make a major contribution to air, soil and water pollution and GHG emissions, and has a profound impact on biodiversity. As such, food systems remain one of the key drivers of climate change and environmental degradation. For this reason, the From Farm to Fork action targets to reduce dependency on pesticides and antimicrobials, reduce excess fertilization, increase organic farming, improve animal welfare, and reverse biodiversity loss.

Biodiversity Strategy for 2030

Biodiversity is also crucial for safeguarding EU and global food security. Biodiversity loss threatens our food systems[20], putting our food security and nutrition at risk. Biodiversity also underpins healthy and nutritious diets and improves rural livelihoods and agricultural productivity[21]. For instance, more than 75% of global food crop types rely on animal pollination.

---

[20] World Economic Forum (2020), The Global Risks Report 2020.

[21] Food and Agriculture Organization (2019), State of the World’s Biodiversity for Food and Agriculture (http://www.fao.org/state-of-biodiversity-for-food- agriculture/en/)

A sustainable food system will be essential to achieve the climate and environmental objectives of the Green Deal, while improving the incomes of primary producers and reinforcing EU's competitiveness.

To contribute to reach the targets of these two programs, the long-term vision of this Chapter includes the following challenges:

- Food security:
  - Intelligent and adaptive food production should take advantage of smart (bio) sensing for high-quality monitoring to reduce the amount of water and chemicals used in such processes, and to prevent contamination.
  - Precision farming systems should require robots with advanced sensing and perception capabilities and drones with intelligent computer vision devices to provide a higher level of detail and on-demand images.
  - Farming Systems should have machine-to-machine interoperable communication (sensors, advanced farming machines and robotic collaborative systems) for cost-effectiveness.
- Food safety:
  - Plants and Animals control; AI should allow to monitor, quantify, and understand individual plants and animals and their variability to control the bio-physical processes (like growing conditions) and understand the biological environment (with plants and animals) to ensure food safety.
  - Plant precision breeding and plant phenotyping should apply large-scale and high-precision measurements of plant growth, architecture, and composition to optimize plant breeding.
  - Integrated pest management should provide smart systems based on portable real-time pest disease diagnostics and monitoring platforms to provide rapid local and regional disease incidence alerts. They should include insect traps.
  - Livestock welfare and health should require smart sensor systems to monitor animal activity to provide useful information for the early detection of diseases and to increase animal wellbeing. They should be also needed for rapid verification of bacterial infection and behavioural observations to control disease spread.
  - Intelligent logistic systems for food chains should require sensing and monitoring of food quality during transport and storage. They should be efficient and interoperable among the logistics chain.
  - End-to-end food traceability should integrate blockchain into current technology to prevent fraud and counterfeiting and provide direct access to end-consumers.
- Environmental protection and sustainable production:
  - In-situ, real-time monitoring of soil nutrients and herbicides should be carried out through intelligent and miniaturized sensors with appropriate packaging. Furthermore, this type of systems should detect weeds, preserve the "good ones" and eradicate the ones that are competing with the crop in question.
  - Air quality monitoring (indoor, urban, and rural) should require the development and

deployment of real-time intelligent multi-sensor technologies with high selectivity and embedded (re-)calibration techniques. Focus should be put in the GHG emission from animals by performing the analysis of the gathered data to support decision making for mitigation of main issues.

  - Smart waste management should provide smart monitoring, controlling waste treatment units in real time as well as gas emissions in landfills and anaerobic digestion monitoring. Data analytics including gamification for behavioural triggers.

- Water resource management:
  - Smart healthy water systems should provide secure drinking water distribution by detecting –in real-time– compounds and contaminants through data analysis capabilities to take the adequate measures to mitigate these issues to secure water quality and its distribution over the network. This requires online information on the status of water sources at larger scales than before. For this, healthy water systems should require connected high-integrated multi-parameter diagnostic sensors for real-time chemical analysis to ensure freshwater.
  - Efficient and intelligent water distribution should require novel smart metering solutions based on various technologies, including electrochemical multi-parameter sensors with high stability, anti-fouling, high accuracy capabilities, and be cost effective. Furthermore, optical sensors based on different principles integrated into miniaturized systems, at a low cost, are also required.
- Biodiversity restoration for Ecosystems Resilience, Conservation and Preservation:
  - Biodiversity restoration for Forestry Ecosystem should provide precision forestry system with remote sensing and AI/ML monitoring capabilities to map and assess the condition of the EU forests as well as early detection and prevention of threats to the forests (wildfires, pests, diseases, etc.). Furthermore, smart systems are required for environment monitoring of forests and fields, as well as CO_2 footprint monitoring. Remotely monitor wildlife behaviour and habitat changes, and provide timely warning upon illegal poaching activity, are also needed.

4.4.6 Digital society

Ubiquitous connectivity ("everywhere and always on(line)") drive people to rely on intelligent applications and the services they use and offer. Public and private infrastructures will increasingly be connected, monitored, and controlled via digital infrastructures ("always measuring") and devices.

Furthermore, the trend of combining of working at the office and from home (or other remote locations), which has been triggered by the Covid-19 pandemic, will continue, and people will endeavour to combine work and private life in other ways.

Digital infrastructures with increased quality of service (QoS) and available bandwidth, will support these trends and will be ubiquitous, both in rural areas as well as in cities. These networks will be open and secure and will support intelligent control management of critical infrastructures, such as water supply, street lighting and traffic. Edge/cloud solutions will arise which will enable increased multimodal situational awareness and ubiquitous localization.

Social inclusion and collective safety and privacy will be enhanced by improved access to public services and communities (as healthcare, education, friends, family, and colleagues), supported by technological innovations in several directions, such as tele-presence, serious gaming, chatbots, virtual reality, robots, and personal and social assistants.

More and more, these solutions will be human-centred, will have cognitive abilities, apply nudging techniques, and support personal development, health, and well-being.

4.5 CONCLUSIONS

The European ECS community, from academia to industry, is a world leader in research, development, and innovation for the past decades. The competition of US and Asian communities is strong and requires significant European effort and investment, so that Europe remains a leader in the coming years, considering the dramatic increase in need for ECS systems due to the emergence of IoT and the corresponding embedded and cyber-physical systems.

In this Chapter, we have presented research and innovation directions for ECS in the long-term, considering the European priorities, such as the Green Deal, and the main objectives of the European ECS community. Considering the interdependence of emerging technologies, application domains and policies that drive innovation, and considering the corresponding trends of European industry in the next few years, we identified long-term challenges for technologies and applications, to provide direction for the community to meet the expected needs of the future. Clearly, the list of challenges and directions that we provide is neither complete nor restricting. Innovation is a continuous process that adapts to new technological capabilities – as they progress–, application needs and, even, new application domains that are not foreseen. However, the current review indicates a clear path to establishing European leadership considering current trends and constraints.

5

Strategic Research and Innovation Agenda 2025

APPENDIX A

Recommendations and Roadmap for European Sovereignty in Open Source Hardware, Software, and RISC-V Technologies

Report from the Open Source Hardware & Software Working Group August 2022

5 APPENDIX A: RECOMMENDATIONS AND ROADMAP FOR EUROPEAN SOVEREIGNTY IN OPEN SOURCE HARDWARE, SOFTWARE, AND RISC-V TECHNOLOGIES

5.1 Executive Summary

The use of open source hardware and software drastically lowers the barrier to design innovative System-On-Chips (SoCs), which is an area that Europe excels in today. Open source is a disruptive approach and strong competences in this field are being developed worldwide, for example in China, the US and India. In order for Europe to become a global player in the field and ensure sovereignty, it is important to put in place a roadmap and initiatives that would permit to consolidate ongoing European activities and to develop new technology solutions for key markets such as automotive, industrial automation, communications, health and aeronautics/defence. Notably open source can be used as a sovereignty tool providing Europe with an alternative to licensing IPs from non-EU third parties. A key success criterion for this is for Europe to develop a fully blown open source **ecosystem** so that a European fork is possible (i.e., create a fully equivalent variant of a given technology), if necessary.

The realization of such an ecosystem requires a radical change in working across the board with leadership and contribution from major European industrial and research players and other value chain actors. An approach similar to the European Processor Initiative which brings together key technology providers and users in the supply chain is needed, but with the goal of producing open source IP. This report suggests the following way forward to foster the development of an open source ecosystem in Europe:

- **Build a critical mass of European open-source hardware/software** that will permit to drive competitiveness, enable greater and more agile innovation, and give greater economic efficiency. At the same time, it will remove reliance on non-EU developed technologies where there are increasing concerns over security and safety. In building this critical mass, it is important to do so strategically by encouraging the design of scalable and re-usable technology.
- **Develop both open source hardware and software as they are interdependent.** An open-source approach to software such as EDA and CAD[2] tools can serve as a catalyst for innovation in open-source hardware. To ensure a thriving ecosystem, it is necessary to have accessible software.
- **Address cross-cutting issues.** In order to enable verticalisation, it is important to address a number of cross-cutting issues such as scalability, certification for safety in different application domains, and security. This requires consideration at both the component level and system level.
- **Cultivate innovation** facilitated through funding from the public sector that is conditional on an open-source approach. The public sector can also have a role in aiding the dissemination of open-source hardware through the deployment of design platforms that would share available IP especially those that were supported through public funds.
- **Engage with the open-source community.** Links with initiatives such as OpenHW Group, CHIPS Alliance, RISC-V International etc. should be strengthened to get and maintain industrial-grade solutions whilst encouraging standardisation where appropriate and keeping links with global open source communities.

To this end, the Working Group that is at the origin of this report has defined a strategic roadmap considering short (2-5 years), medium (5-10 years) and long term (>10 years) goals. The success of the roadmap depends on European actors working closely together to create a critical mass of activities that enhance and expand the European open source community and its influence on the world stage. The Working Group advocates that this roadmap of activities is supported via coordinated European level actions to avoid fragmentation and ensure that Europe retains technological sovereignty in key sectors.

5.2 Introduction

Europe has a core competence in designing innovative System-On-Chips (SoC) for key application sectors such as automotive, industrial automation, communications, health and defense. The increasing adoption of open source, however, is potentially disruptive as it drastically lowers the barrier to design and to innovation. Already nations such as China, the US and India are investing heavily in open source HW and SW to remain competitive and maintain sovereignty in key sectors where there are increasing concerns over security and safety.

Europe needs to respond by creating a critical mass in open source. The development of a strong European open source ecosystem will also drive competitiveness as it enables greater and more agile innovation at much lower cost. However, to achieve this there is a need to align and coordinate activities to bring key European actors together.

This document proposes a high-level roadmap for Open Source HW/SW and RISC-V based Intellectual Property (IP) blocks which are of common interest to many European companies and organizations, considering design IP and supporting EDA tools. The aim of the roadmap is to consolidate ongoing European activities through the presentation of a number of concrete ideas, topics and priority areas for IP-blocks and future research in order to build sovereignty and competitive advantage in Europe towards future energy-efficient processor design and usage.

Chapter 2 presents an initial approach to open source, shows how sovereignty can be addressed via open source, and what Europe needs to do to create a strong and competitive critical mass in this area. This is followed by a roadmap for open-source hardware in Europe in Chapter 3, which also highlights the necessary toolbox of open-source IP for non-proprietary hardware to thrive. Chapter 4 presents the supporting software required for the ecosystem to flourish. Chapter 5 discusses some cross-cutting issues – such as scalability, safety and security – for a potential RISC-V processor in common use. Chapter 6 identifies gaps in the current European ecosystem for both hardware and software. In Chapter 7, some important elements of the proposed roadmap are discussed. In turn, Chapter 8 lists the short-, mid- and long-term needs for the different elements of the proposed roadmap. Chapter 9 presents a series of recommendations of both a global and specific nature. Chapter 10 lists a series of horizontal and vertical activities to address the needs of open-source development within Europe. Some concluding remarks and two annexes complete the report.

5.3 Views of an Open Source Strategy

5.3.1 Introduction

The interest in open source is rapidly rising as shown in **Figure 5.1** with the number of published academic papers exploding in the area of RISC-V. This presents a major discontinuity for SoC design. The use of open source drastically lowers the barrier to design innovative SoCs which is an area that Europe excels in and strongly depends on in key application sectors. Use of open source also allows a research center or company to focus its R&D effort on innovation, leveraging an ecosystem of pre-validated IP that can be freely assembled and modified. This is disruptive in a market where traditionally a set of IPs is designed in house at a cost that is only accessible to a few companies, or where alternatively 3rd party IPs are licensed by companies with constraints on innovation due to architecture. A key advantage of open source is that it provides a framework that allows academia and companies to cooperate seamlessly, leading to much faster industrialization of research work.

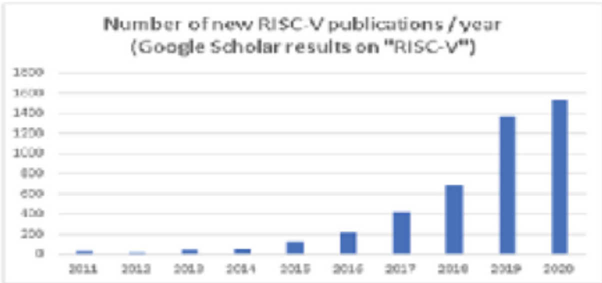


Figure 5.1 - Google Scholar results on RISC-V (search performed 1 Sept 2021) showing growth of annual publications persisting even despite COVID-19 conference cancellations

There are many parallels between open source hardware and open source software. There have been key exemplars in the software domain such as GNU and Linux. The latter started as an educational project in 1991 and has now become dominant in several verticals (High Performance Computing (HPC), embedded computing, servers, etc.) in synergy with open source infrastructure such as TCP/IP network stacks.

While there are a lot of commonalities in term of benefits, pitfalls and the transformative potential of the industry, there are also significant differences between open source hardware and open source software coming from the physical embodiment of any Integrated Circuit (IC) design via a costly and timely process.

The ecosystem is wide ranging and diverse including the semiconductor industry, verticals and system integrators, SMEs, service providers, CAD tools providers, open source communities, academics and research. It is an area with many opportunities to create innovative start-ups and service offers. The benefits and attraction of adoption of open source depends on the type of actor and their role within the value chain, and can include creating innovative products with lower costs and access barriers, and providing a faster path to innovation.

Annex A gives more details on open source, including potential benefits, the open source ecosystem, key players, business strategies, licensing models, and licensing approaches.

5.3.2 Current and growing market for RISC-V

Annex B presents an overview of the market trends in the global chips market. The value of this market in 2021 was around USD $550 billion and is expected to grow to USD $1 trillion in 2030.

RISC-based solutions are being used in a growing range of applications and the uptake of RISC-V is expected to grow rapidly.

Within these expanding markets RISC-based architectures have become important. RISC platforms have long enabled mobile phone and small-device consumer electronics vendors to deliver attractive price, performance and energy use compared to complex instruction set computing (CISC)-based solutions such as those of Intel and AMD. Also increasingly, RISC-based solutions are being used to power servers, moving beyond their small-device origins[3] as shown in **Figure 5.2** with a rapid growth in the uptake of RISC-V[4].

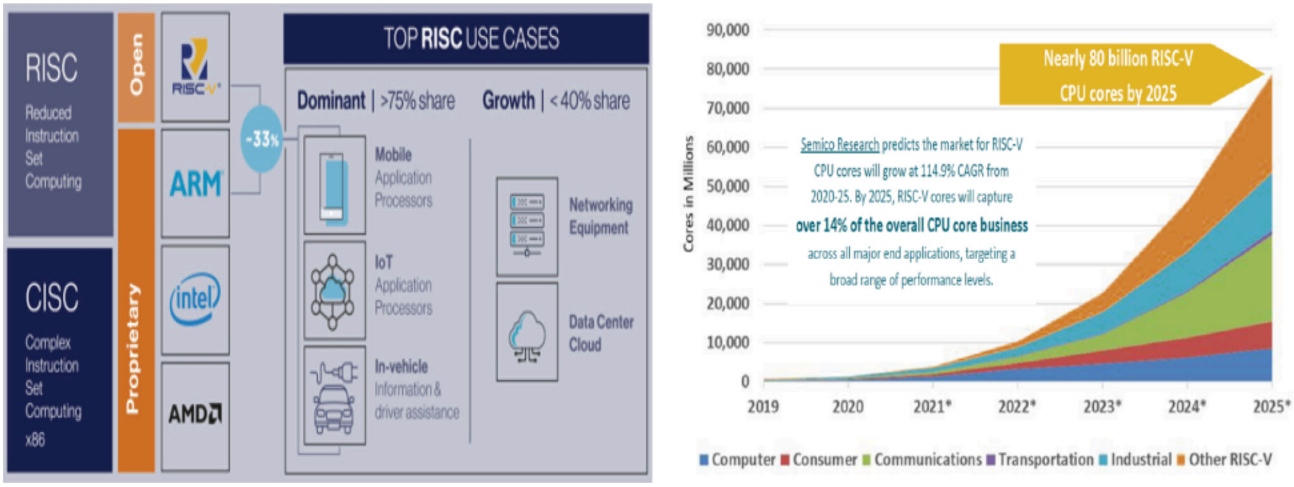


Figure 5.2 - Growing RISC-based Architectures Usage (Source: Altman Solon and Semico Research Corp.)

The compound annual growth rate for RISC-V from 2018-2025 is given in **Figure 5.2**[5]. In key European markets, the RISC-V penetration rate is expected to grow rapidly[6]. According to "Wilson Research Group Functional Verification" **in 2020**[7]**, 23% of projects in both the ASIC and FPGA spaces incorporated at least one RISC-V processor**.

The current RISC-V instruction set is very popular with RTOs and academics and is also attracting vertical companies that want to develop their own in-house semiconductors (Western Digital[8], Alibaba[9], ...). This has led to a number of start-up companies adopting RISC-V to develop specific processors or processor families to sell to the market. Some European examples of success stories, are:

- Greenwaves Technologies (founded in 2014): ultra-low power and AI-enabled GAP processor family for energy constrained products such as hearable, wearables, IoT and medical monitoring products.
- Codasip (founded in 2014): their processor core portfolio combines two complementary processor families, Codasip RISC-V Processors and SweRV Cores, to cover a wide range of design needs. Their business consists in selling of the standard cores that they have developed, as well as to customize the cores to the needs of the customer.
- Cobham Gaisler (founded in 2001): NOEL-V processor family based on 32-bit and 64-bit RISC-V including fault-tolerant implementation and Hypervisor extension, focusing on application of processors in harsh environments.

5.3.3 European Strategy

Open source shows lots of promise to boost the economy and welfare in our societies, but this is a non-trivial task. As illustrated with the RISC-V example, a generally agreed mechanism is to use open source to create or enlarge markets and to use proximity, prime-mover and other effects to benefit preferentially from those new markets, combining mostly publicly funded open source and privately funded proprietary offerings. This is well known already for software, and the aim is to replicate this for hardware, allowing the research and industry community to focus on new innovative points, rather than on re-developing a complex framework such as a processor and its ecosystem (compiler, interfaces, ...). There is no way to "protect IP" while still building real open source other than by providing adequate financing for Europe-based entities who work with open source as well as support for regional initiatives and leadership in identifying practical problems that can be solved. There is also a need to help connect open source vendors with customers in need of the innovation. **Speed from concept to product is also key to stay ahead of non-European competitors. This requires foundries and EU financing to build an ecosystem that can get from concept to products quickly in order to maintain European sovereignty.**

Europe is starting initiatives in terms of public adoption of Linux or Matrix in governmental institutions and other areas, and hardware presents a unique opportunity where the ecosystem is immature enough that it could really make a difference. Education is a key factor which could help. Talented and educated Europeans tend to stay in Europe, so significantly boosting open source hardware education could be a way to seed a strong ecosystem. RISC-V is already part of the curriculum in many universities and **Europe could further fund open source hardware education at an even earlier level and generally strongly encourage universities to produce open source.**

The success of an open source project requires efforts going beyond putting some source code in a public repository. The project needs to be attractive and give confidence to its potential users. It is said that the first 15 minutes are the most important. If a potential user cannot easily install the software (in case of open source software), or does not understand the structure, the code and the parameters, it is unlikely that the project will be used. Having good documentation and well written code is important. This requires time and effort to clean the code before it is made accessible. It is also necessary to support the code during its lifetime which requires a team to quickly answer comments or bug reports from users and provide updates and new releases. All modifications should be thoroughly checked and verified, preferably via automated tools. Without this level of support potential users will have little trust in the project. There is thus a need for support personnel and computing resources. This is a challenge for European Open Source projects as the code needs to be maintained after the official end of the project. Only if a project is attractive, provides innovation, is useful and is trusted, will it attract more and more users and backers who can contribute, correct bugs, etc. At this point it can be self-supported by the community with less effort from its initial contributors. The efforts and resources required to support an open source project should not be under-estimated. The project also needs to be easily accessible and hosted in a reliable industrial environment such as the Eclipse foundation, where the Open Hardware Group can provide support for ensuring the success of an open source project.

Success of open source depends not only on the IP blocks but also on the documentation, support and maintenance provided. This requires significant resources and personnel.

It is important to help steer Europe towards open source hardware and make sure that knowledge is shared, accessible, maintained and supported on a long-term basis. Barriers to collaboration need to be eliminated. Priorities should be discussed considering what is happening elsewhere in the world and **to maximize the value generated within Europe there is a need to build value chains, develop a good implementation plan and provide long-term maintenance and support**.

A more detailed economic view on the impact of Open Source HW and SW on technological independence, competitiveness and innovation in Europe has been documented in a report edited by Fraunhofer ISI and OpenForum Europe[17].

5.3.4 Maintaining Europe's Sovereignty and Competitiveness

A key market challenge is that most of the current open source hardware technology resources are located outside of Europe, especially in the United States, in China (which has strongly endorsed RISC-V based open source cores) and India which has a national project. This raises the question: how can Europe maintain sovereignty and stay competitive in a rapidly developing open source market? The traditional value of the European landscape lies in its diversity and collaborative nature which reflects the nature of open source very well. EU funding looks favorably on open source solutions and there are well-established educational institutions which are of a less litigious / patent-heavy nature as compared to US private universities. The diversified industry within Europe is also a strength that can be leveraged to push the EU's competitiveness in open hardware.

To get sufficient return on investment for Europe, a focus should be given to application domains where there is stronger impact: automotive, industrial automation, communications, health, defense, critical systems, IoT, cybersecurity, etc. These application domains convey specific requirements towards open source technologies: safety, security, reliability, power and communication efficiency.

A key to sovereignty in the context of open source is the involvement of sufficient European actors in the governance of the various projects (CHIPS alliance, OpenHW group, etc.) and the achievement of a critical mass of EU contributors to these projects so that a fork could be pursued if this is forced onto EU contributing members.

Competitiveness in the context of open source should be looked at in a reversed manner, that is, how much competitiveness European players would lose by not adopting open source. Open source is becoming a major contributor to innovation and economic efficiency of the players who broadly adopt the approach. The massive on-going adoption of open source in China, from leading companies to start-ups and public research centers, with strong support from both the central and regional authorities, is a very interesting trend in China's strategy to catch up in semiconductors.

Recommendation – Calls by the Key Digital Technologies Joint Undertaking (KDT JU) should bring benefits to open hardware. Use of this hardware by proprietary demonstrations (including software, other hardware, mechanical systems, in key sectors automotive, industrial automation, etc.) would be beneficial for pushing acceptance of open hardware. Calls could formulate a maximum rate (e.g. 30% of volume) for such demonstrations. It should be required that proposals do not merely implement open interfaces (e.g. RISC-V) but also release implementations (e.g. at least 50% of volume). It is reasonable to demand that the open source licenses of such released implementations allow combination with proprietary blocks.

While this document focuses on open source IC design IPs, the major control point from a sovereignty standpoint is EDA (Electronic Design Automation) tools, which are strictly controlled by the US, even for process nodes for which Europe has sovereign foundries. Europe has significant capabilities in this area in research, which it regularly translates into start-ups for whom the only exit path currently is to be acquired by US controlled firms. Open source EDA suites exist but are limited to process nodes for which PDK is considered as a commodity (90nm and above). **A possible compromise approach, complementary to the development of open source EDA, is to emphasize compatibility of open source IP design with proprietary, closed source design tools and flows**. This objective can be achieved with efforts on two fronts: (I) promote licensing agreement templates for commercial EDA tools that explicitly allow the development of open hardware: this is especially important when open hardware is developed by academic partners in the context of EU-funded R&D actions, as current academic licenses (e.g., as negotiated by EUROPRACTICE) can often not be used for commercial industrial activities. And (II) even more importantly, emphasis should be put in promoting open standards for data exchange of input and outputs of commercial EDA tools (e.g., gate-level netlists, technology libraries). **Proprietary specifications and file exchange formats can hinder the use of industry-strength tools for the development of open source hardware**. This matter requires a specific in-depth analysis to propose an actionable solution.

5.3.5 Develop European Capability to Fork if Needed

Open source also brings more protection against export control restrictions and trade wars, as was illustrated in 2019 when the US Administration banned Huawei from integrating Google proprietary apps in their Android devices. However, Huawei was not banned from integrating Android as it is open source released. Regarding US export rules, a thorough analysis is provided by the Linux Foundation[18].

Even for technologies developed in other parts of the world that have been open sourced and adopted in Europe at a later stage, **the ability to potentially "fork" i.e., create a fully equivalent variant of a given technology is a critical capability from a digital sovereignty perspective**. Should another geopolitical block decide to disrupt open source sharing by preventing the use of future open source IPs by EU players, the EU would need to carry on those developments with EU contributors only, or at least without the contribution of the adversary block. **For this there needs to be a critical mass of European contributors available to take on the job if necessary.**

The realization of such a critical mass will require an across-the-board change of working, with leadership and contribution from major European players (industrial and research) as well as a myriad other contributors. To achieve this there is a need to build or take part in sustainable open source communities (OpenHW Group, CHIPS Alliance, etc.) to get and maintain industrial-grade solutions. Care needs to be taken to avoid fragmentation (creating too many communities) or purely European communities with a disconnection from global innovation. A challenge is that current communities are young and essentially deliver processor cores and related SW toolchains. They need to extend their offer to richer catalogues including high-end processors, interconnects, peripherals, accelerators, Operating Systems and SW stacks, IDE and SDK, extensive documentation, etc.

Open source hardware targeting ASIC implementations requires software tools for implementation where licensing costs for a single implementation quickly amount to several hundred thousand Euro. **There is thus a need for high quality open source EDA tools supporting industrial-grade open source IP cores**. Europe also has a low footprint in the world of CAD tools, which are critical assets to design and deliver electronics solutions. Recent open source CAD tool initiatives are opportunities to bridge this gap.

Considering processors **involvement in RISC-V International should be encouraged** to influence and comply with the "root" specifications. RISC-V standardizes the instruction set, but additional fields will be needed in future for interoperability: SoC interconnect, chiplet interfaces, turnkey trust and cybersecurity, safety primitives, etc. **Europe should promote open source initiatives that would help maximize collaborations and reach a significant European critical mass. This is a key issue if Europe is to compete with China and the USA**. **A significant portion of the intellectual property produced in these European initiatives should be delivered as open source**, so that European actors can exploit these results.

	Computer	Consumer	Communications	Transportation	Industrial	Other RISC-V	Total
Adv. Perf Multicore SoC	60%	73%	224%	153%	106%	171%	144%
Value Multicore SoC	60%	92%	222%	159%	122%	201%	177%
Basic SoC	63%	115%	217%	166%	127%	215%	190%
FPGA	62%	72%	190%	163%	102%	176%	149%
Total	61%	81%	209%	160%	110%	185%	158%

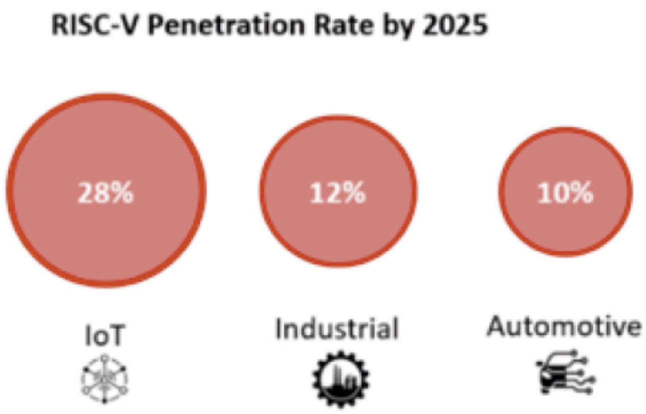


Figure 5.3 -CAGR for RISC-V and Predicted Penetration Rates in key European sectors

5.4 Towards an Open Source Hardware and Software Roadmap for Europe

5.4.1 Processors (RISC-V, beyond RISC-V, ultra-low power and high-end)

The Working Group has identified the strategic key needs for the development and support for:

1. a range of different domain focused processors,

2. the IP required to build complete SoCs and,

3. the corresponding software ecosystem(s) for both digital design tools and software development.

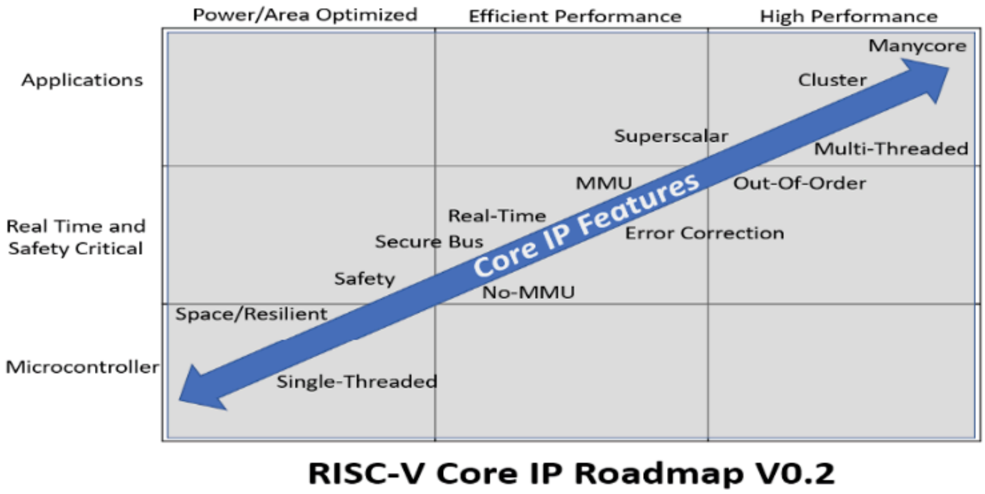


Figure 5.4 -RISC-V Core IP Roadmap V0.2

To align the software roadmap to the hardware IP core development roadmap, **efforts should be focused on supporting RISC-V implementations that correspond to the RV64GC and RV32GC set of extensions**. At the same time the creation/adoption of RISC-V profiles (microprocessor configurations) should be encouraged in the community. To scope this there is a need for an overall roadmap for RISC-V microprocessor core IP for specific market focus areas. A challenge is that microprocessor core IP comes in many shapes and sizes with differing requirements to meet the needs of applications ranging from 'toasters to supercomputers' as shown in **Figure 5.4**. In order to pursue this the Working Group has concentrated on a subset of this broader landscape whilst leveraging the underlying logic and thought process.

To be successful the IP roadmap must encompass attributes such as performance, scalability, software compatibility and deliver high levels of architectural reuse over many years of robust product creation. The focus has been on microprocessor core IP that drives system solutions ranging from edge IOT devices to manycore compute platforms which are key to many European applications. This needs to meet the challenges of scalability and must be sustainable in years to come. At the same time software and digital design tooling needs to be developed to support the foundational RISC-V IP. A key aim has been to leverage global open source hardware and software projects to maximize European reuse, avoid duplication and strategically direct investment.

5.4.2 Accelerators – Domain Specific Architectures

Accelerators are a key need and for many application domains, as meeting non-functional requirements (e.g., performance, energy efficiency, etc.) is often difficult or impossible using general-purpose processor instruction sets such as RV32IM or RV64GC. This is true, for example, in machine learning and cryptography, but it also applies to high-performance storage and communications applications. The use of extended instruction sets to enable more parallel processing, such as RISC-V “P” for Single Input Multiple Data (SIMD) processing and RISC-V “V” for full-scale vector processing, can provide significant performance and/or efficiency gains, however, even higher gains are achievable in many cases by adding application domain-specific features to a hardware architecture.

This has led to the concept of Domain-Specific Architectures (DSA) which was highlighted by computer architecture pioneers John Hennessy and David Patterson in their 2018 Turing Award Lecture[19] (the “Nobel Prize” of computer science and computer engineering). Here Domain-Specific Architectures were noted as one of the major opportunities for the further advances in computer architecture to meet future non-functional targets.

Accelerators are one approach to implement a DSA. They do not provide all of a system’s functionality, but instead assist more general-purpose processors by the improved execution of selected critical functions. In many cases, this “improvement” means “faster” execution, leading to the name “accelerator”, however, it can also mean “more energy efficient” or “more secure”, depending on the requirements. In order for the accelerator to exceed the capabilities of a general-purpose system on a similar semiconductor process node, the micro-architecture of the processor is highly specialized and very specific to the actual algorithm it implements. For example, machine-learning accelerators dealing with the efficient inference for Artificial Neural Networks will have very different micro-architectures depending on whether they aim to operate on dense or sparse networks. They, in turn, will have very different architectures from accelerators dealing, e.g., with 3D-Stereovision computations in computer vision for autonomous driving, or Quantum Computing-resistant cryptography for secure communications.

In most cases, accelerators rely on conventional digital logic and are designed/implemented/verified as is usual for the target technology (e.g., FPGA, ASIC). They would thus profit immediately from all open source advances in these areas, e.g., open source EDA tools. Open sourcing the accelerator designs themselves will benefit the ecosystem by encouraging reuse, standardization of testing and verification procedures and in inducing more academic interest and research. Some open sourced accelerators have seen widespread use. Successful cases include NVIDIA’s NVDLA[20] Machine Learning inference accelerator, or the 100G hardware-accelerated network stack from ETHZ[21].

The opportunity for much higher impact when using open source for accelerators lies not so much in the accelerators themselves, but in the

hardware and software interfaces and infrastructures enabling their use.

Accelerators are integrated into the surrounding computing system typically in one of three approaches (although others exist, e.g., network-attached accelerator appliance).

Custom Instructions: At the lowest level, the accelerator functions can be tightly integrated with an existing processor pipeline and are accessed by issuing special instructions, sometimes called custom instructions or ISA eXtensions (ISAX). This model is suitable for accelerator functions that require/produce only a limited amount of data when operating and thus easily source their operands and sink their result from/to one of the processor's existing register files (e.g., general-purpose, floating-point or SIMD). The key benefit of this approach is the generally very low overhead when interacting with the accelerator from software. This approach has been exploited for decades in processors such as the proprietary Tensilica and ARC cores and is also used, for example, in Intel's AES-NI instructions for accelerating Advanced Encryption Standard (AES) cryptography. In the RISC-V domain, companies such as Andes Technologies (Taiwan) emphasize the easy extensibility of their cores with custom functionality. In the open source area, ETHZ has implemented their Xpulp[22] instructions, showing significant performance gains and code-size reduction over the base RISC-V ISA.

Custom instructions generally require deep integration into a processor's base microarchitecture that is often difficult with many proprietary offerings, e.g., from Arm. In the open source RISC-V hardware ecosystem, though, custom instructions have become a very effective means towards Hennessy & Patterson's Domain-Specific Architectures. **Examples of custom functionality added in this manner include Finite-Field Arithmetic for Post-Quantum Cryptography [23], digital signal processing [24], or machine-learning[25].**

However, due to their need for deep integration into the base processor, such custom instructions are generally not portable between cores. What would be desirable is a lightweight, flexible interface for the interactions between base core and the accelerator logic realizing the custom functionality. This interface would need to be bidirectional as for example instruction decoding would still be performed by the base core, and only selected fields of the instruction would be communicated to the accelerator. The accelerator, in turn, would pass a result to the base core for write-back into the general-purpose register file. Additional functionality required includes access to the program counter computations, to realize custom control flow instructions (this is missing from UCB's RoCC interface), and the load/store-unit(s) for custom memory instructions. A key aspect of an efficient custom instruction interface will be that the unavoidable overhead (hardware area, delay, energy) is only paid for those features that are actually used. For example, a simple Arithmetic Logic Unit (ALU) compute operation should not require interaction with the program counter or memory access logic**. With a standard (de-facto or formal) interface in place, R&D work on tightly integrated accelerators could port across different base processors.**

Loosely Coupled Accelerators: More complex accelerators that require or produce more data than can easily be provided/accepted by a base processor core are generally not integrated into the core pipeline itself but are coupled more loosely using industry-standard protocols such as Arm AMBA. The communication requirements of the accelerator determine the specific flavour of interface used, which can range from Advanced Peripheral Bus (APB) for low-bandwidth interaction, or Advanced High-Performance Bus (AHB)/Advanced eXtensible Interface (AXI) for higher-performance accelerators. These loosely coupled accelerators accept commands from one or more base cores using memory-mapped interactions and are then capable of accessing their input and output data independently using Direct Memory Access (DMA) operations. Since these accelerators are integrated into a system using standard protocols, they are very portable. **For example, the open source NVIDIA machine-learning accelerator has been successfully integrated into a number of open source and proprietary Systems-on-Chips.**

However, for more complex accelerators, the open source ecosystem is much sparser when advanced capabilities are required. Two examples of this are shared-virtual memory and/or cache-coherent interactions between the accelerators and the software-programmable base cores. Both mechanisms can considerably simplify using the interaction between software on the base cores and accelerators, especially when more complex data structures (e.g., as used in databases) need to be operated on. To achieve shared-virtual memory, the accelerator and the base core(s) need to share address translations, invalidations, and fault-handling capabilities. The UCB RoCC interface achieves this by leveraging the capabilities of the underlying base core (Rocket, in most cases). However, for better scalability across multiple accelerators, it would be beneficial to have a dedicated Input-Output Memory Management Unit (IOMMU) serving just the accelerators, possibly supported by multiple accelerator-specific Translation-Lookaside Buffers (TLBs) to reduce contention for shared resources even further. However, no such infrastructure exists in an open source fashion. Currently, designers following Hennessy & Patterson's ideas towards accelerator-intensive systems do not just have to design the (potentially complex) accelerators themselves, but they also often have to start "from scratch" when implementing the supporting system-on-chip architecture allowing these accelerators to actually operate. **Portable, scalable and easily accessible open source solutions are sorely needed to lower this barrier of entry.**

Recommendation - Low-level generic and higher-level domain-specific frameworks should be made available in an easily accessible open source manner. For the lower-level frameworks, the goal should be portable and scalable solutions that support the selected model(s) of computation in a lightweight modular manner.

High-Speed DRAM Interfaces: Accelerators are often employed for highly data-intensive problems (e.g., graphics, vision, machine-learning, cryptography) that need to store significant amounts of data themselves, and/or require high-performance access to data via a network. Thus, the availability of high-speed interfaces to Dynamic Random Access Memory (DRAM), ideally even in the form of forward-looking 2.5D/3D memory technologies such as High Bandwidth Memory (HBM), and to a fast network or peripheral busses such as PCI Express, are absolutely crucial for the use of and research into accelerators. In most cases, it does not make sense to design an accelerator if it cannot interact with sufficient amounts of data. Moving up further in the layers of systems architecture, accelerators are not just used today as IP blocks in a system-on-chip, but also as discrete expansion boards added to conventional servers, e.g., for datacenter settings.

Recommendation - To enable easier design and use of novel accelerator designs open source system-on-chip templates should be developed providing all the required external (e.g., PCIe, network, memory) and internal interfaces (e.g., Arm AMBA AXI, CHI, DRAM and interrupt controllers etc.), and into which innovative new accelerator architectures can be easily integrated. These template SoCs could then be fitted to template PCBs providing a suitable mix of IO and memory (as inspired by FPGA development boards) to the custom SoC. Ideally, all of this would be offered in a "one-stop-shop" like approach, similar to the academic/SME ASIC tape-outs to EUROPRACTICE for fabrication but in this case functional PCIe expansion boards.

This not only applies to Graphics Processing Units (GPUs), as the most common discrete accelerator today, but also to many machine learning accelerators such as Google's TPU series of devices. This usage mode will become even more common, now that there is progress on the required peripheral interfaces for attaching such boards to a server, specifically: Peripheral Component Interconnect Express (PCIe) has finally picked up again with PCIe Gen4 and Gen5. Not only do these newer versions have higher transfer speeds, they also support the shared-virtual memory and cache-coherent operations between host and the accelerator board, using protocol variants such as Cache Coherent Interconnect for Accelerators (CCIX) or Compute Express Link™ (CXL). **Designing and implementing a base board for a PCIe-based accelerator, though, is an extremely complex endeavor.** Not just from a systems architecture perspective, but there are also many practical issues, such as dealing with high-frequency signals (and their associated noise and transmission artifacts), providing cooling in the tight space of a server enclosure, and ensuring reliable multi-rail on-board power-supplies.

These issues are much simplified when using one of the many FPGA based prototyping/development boards which have (mostly) solved these problems for the user. High-speed on-chip interfaces are provided by the FPGA vendor as IP blocks, and all of the board-level hardware comes pre-implemented. While these boards are generally not a perfect match to the needs of a specific accelerator (e.g., in terms of the best mix of network ports and memory banks), a reasonable compromise can generally be made choosing from the wide selection of boards provided by the FPGA manufacturers and third-party vendors. This is not available, however, to academic researchers or SMEs that would like to demonstrate their own ASICs as PCIe-attached accelerators. **To lower the barrier from idea to usable system in an open source ecosystem, it would be highly desirable if template Printed Circuit Board (PCB) designs like FPGA development boards were easily available, into which accelerator ASICs could easily be inserted, with all of the electrical and integration issues already being taken care of.** Providing these templates with well-documented and verified designs will incentivize designers to release their work in the public domain.

Hardware is one aspect of an accelerator, but it also needs to be supported with good software. This can range from generic frameworks, e.g., wrapping task-based accelerator operations using the TaPaSCo[26] system, to domain-specific software stacks like Tensor Virtual Machine (TVM [27] for targeting arbitrary machine-learning (ML) accelerators. Combinations of software frameworks are also possible, e.g., using TaPaSCo to launch and orchestrate inference jobs across multiple TVM-programmed ML accelerators [28].

Recommendatio**n** - For higher-level frameworks, future open source efforts should be applied to leveraging existing domain-specific solutions. Here, a focus should be on making the existing systems more accessible for developers of new hardware accelerators. Such an effort should not just include the creation of "cookbook"-like documentation, but also scalable code examples that can be easily customized by accelerator developers, without having to invest many person-years of effort to build up the in-house expertise to work with these frameworks.

To enable broader and easier adoption of accelerator-based computing, a two-pronged approach would be most beneficial. Lower-level frameworks, like TaPaSCo, should provide broad support for accelerators operating in different models of computation, e.g., tasks, streams/dataflow, hybrid, Partitioned Global Address Space **(**PGAS), etc., drawing from the extensive prior work in both theoretical and practical computing fields. Many ideas originating in the scientific and high-performance computing fields originally intended for supercomputer-scale architectures have become increasingly applicable to the parallel system-on-chip domain. Lower-level frameworks can be used to provide abstractions to hide the actual hardware-level interactions with accelerators, such as programming control registers, setting up memory maps or copies, synchronizing accelerator and software execution, etc. Application code can then access the accelerator using a high-level, but still domain-independent model of computation, e.g., launching tasks, or settings up streams of data to be processed by the hardware.

Even higher levels of abstraction, with their associated productivity gains, can be achieved by making the newly designed accelerators available from domain-specific frameworks. Examples include the TVM for machine learning, or Halide[29] for high-performance image and array processing applications. These systems use stacks of specialized Intermediate Representations (IR) to translate from domain-specific abstractions down to the actual accelerator hardware operations (e.g., TVM employs Relay[30], while Halide can use FROST[31]). Ideally, to make a new accelerator usable in one of the supported domains would just require development of a framework back-end for mapping from the abstract IR operations to the accelerator operations and the provision of an appropriate cost-model, to allow the automatic optimization passes included in many of these domain-specific frameworks to perform their work.

It should be noted that the development of OpenCL 2.x is a cautionary tale of what to avoid **when designing/implementing such a lower-level framework**. Due to massive overengineering and design-by-committee, it carried so many mandatory features as baggage that most actual OpenCL implementations remained at the far more limited OpenCL 1.x level of capabilities. This unfortunate design direction, which held back the adoption of OpenCL as a portable programming abstraction for accelerators for many years, was only corrected with OpenCL 3.0. This version contains a tightly focused core of mandatory functionality (based on OpenCL 1.2) supported by optional features that allow tailoring to the specific application area and target accelerators.

From a research perspective, it would also be promising to study how automatic tools could help to bridge the gap between the domain-specific frameworks, e.g., at the IR level, and the concepts used at the accelerator hardware architecture levels. Here, technologies such as the Multi-Level IR (MLIR) proposed for integration into the open source Low Level Virtual Machine (LLVM[32]) compiler framework may be a suitable starting point for automation.

5.4.3 Peripherals and SoC Infrastructure

5.4.3.1 SoC Infrastructure

In addition to processor cores, it is also very important to have a complete infrastructure to make SoCs and be sure that all IPs are interoperable and well documented (industry grade IPs). This requires the necessary views of IPs (IDcard with maturity level, golden model, Register-Transfer Level (RTL), verification suite, integration manual, Design For Test (DFT) guidelines, drivers) which are necessary to convince people to use the IPs. As highlighted system-on-chip (SoC) templates are needed that provide all the required external (e.g., PCIe, network, memory) and internal interfaces and infrastructures (e.g., Arm AMBA AXI, Coherent Hub Interface (CHI), DRAM and interrupt controllers, etc.), and into which innovative new IPs could be easily integrated. High-speed lower-level physical interfaces (PHYs) to memories and network ports, are designed at the analog level, and are thus tailored to a specific chip fabrication process. The technical details required to design hardware at this level are generally only available under very strict NDAs from the silicon foundries or their Physical Design Kit (PDK) partners. Providing access to this information would involve inducing a major shift in the industry. As a compromise solution, though, innovation in open source hardware could profit immensely if these lower-level interface blocks could be made available in a low-cost manner, at least for selected chip fabrication processes supported by facilitators such as EUROPRACTICE for low-barrier prototyping (e.g., the miniASIC and MPW programs).

Recommendation - Blocks do not solely consist of (relatively portable) digital logic, but they also have analog components in their high-speed physical interfaces (PHY layer) that must be tailored to the specific ASIC fabrication process used. Such blocks should be provided for the most relevant of the currently EUROPRACTICE-supported processes for academic/SME prototype ASIC runs. The easy exchange of IP blocks between innovators is essential leveraged by both the EDA companies and IC foundries. An exemplar is the EUROPRACTICE enabled R&D structure of CERN where different academic institutions collaborate under the same framework to exchange IP.

For open source hardware to succeed, standard interfaces such as DRAM controllers, Ethernet Media Access Controllers (MACs) and PCIe controllers need to be either available as open source itself, or at a very low cost at least for academic research and SME use. If they are not, then the innovation potential for Europe in both of these areas will often go to waste, because new ideas simply cannot be realized and evaluated beyond simulation in real demonstrators/prototypes. For a complete SoC infrastructure, it is necessary to have an agreed common interconnect at the:

- processor/cache/accelerator level (similar to the Arm Corelink Interconnect, or the already existing interconnect specifications for RISC-V such as TileLink.[33]): in the open source area, support for cache-coherency exists in the form of the TileLink protocol, for which an open source implementation is available from UC Berkeley. However, industry standard protocols such as AMBA ACE/CHI do not have open source implementations (even though their licenses permit it). This makes integration of existing IP blocks that use these standard interfaces with open source systems-on-chips difficult.
- memory hierarchy level, for example between cores for many-core architectures (NUMA effects) and between cores and accelerators,
- peripheral level (e.g., Arm's AMBA set of protocols).

The communication requirements of the accelerators determine the specific flavor of interface used, which can range from APB for low-bandwidth interaction, or AHB/AXI for higher-performance accelerators.

5.4.3.2 Networks on a Chip

Networks on a Chip (NoCs) and their associated routers are also important elements for interconnecting IPs in a scalable way. Depending on requirements, they can be synchronous, asynchronous (for better energy management) or even support 3D routing. Continuous innovations are still possible (and required) in these fields.

5.4.3.3Verification and Metrics

IPs need to be delivered with a verification suite and maintained constantly to keep up with errata from the field. **For an end-user of IP the availability of standardized metrics is crucial as the application scenario may demand certain boundaries on power, performance, or area of the IP.** This will require searches across different repositories with standardized metrics. One industry standard benchmark is from the Embedded Microprocessor Benchmark Consortium (EEMBC)[34], however, the topic of metrics does not stop at the typical performance indicators. It is also crucial to assess the quality of the verification of the IP with some metrics. In a safety or security context, it is important for the end-user to assess in a standardized way the verification status in order to conclude what is still needed to meet required certifications. **Using standardized metrics allows end users to pick the most suited IP for their application and get an idea on needed additional efforts in terms of certifications.**

Another aspect is in providing trustworthy electronics. This is a continuous effort in R&D, deployment and operations, and along the supply chains. This starts with trustworthy design IPs developed according to auditable and certifiable development processes, which give high verification and certification assurance (safety and/or security) for these IPs. These design IPs including all artefacts (e.g., source code, documentation, verification suites) are made available ensuring integrity, authenticity and traceability using certificate-based technologies. Traceability along the supply chain of R&D processes is a foundation for later traceability of supply chains for components in manufacturing and deployment/operation.

5.4.3.4 Chiplet and Interposer Approach

Another domain which is emerging and where Europe can differentiate itself is in using the 2.5D approach, or "chiplets + interposers". This is already enabled by EUROPRACTICE for European academics and SMEs. **The idea is to assemble functional circuit blocs (called chiplets, see https://en.wikichip.org/wiki/chiplet) with different functions (processor, accelerator, memories, interfaces, etc.) on an "interposer" to form a more complex System-on-Chip**. The interposer technology physically hosts the chiplets and ensures their interconnection in a modular and scalable way, like discrete components on a Printed Circuit Board. This approach can range from low-cost systems (with organic interposers), to high-end silicon-based passive and active interposers, up to photonic interposers. In active interposers, the interposer also includes some active components that can help with interconnection (routers of a NoC), interfacing or powering (e.g., Voltage converters).

The industry has started shifting to chiplet-based design whereby a single chip is broken down into multiple smaller chiplets and then "re-assembled" thanks to advanced packaging solutions. **TSMC indicates that the use of chiplets will be one of the most important developments in the next 10 to 20 years. Chiplets are now used by Intel, AMD and Nvidia and the economics of this has already been proved by the success of the AMD chiplet-based Zen architecture.** As shown by the International Roadmap for Devices and Semiconductors (IRDS) roadmap, chiplet-based design is considered as a complementary route for More Moore scaling.

Chiplet-based design is an opportunity for the semiconductor industry, but it creates new technical challenges along the design value chain: architecture partitioning, chiplet-to-chiplet interconnect, testability, CAD tools and flows and advanced packaging technologies. None of the technical challenges are insurmountable, and most of them have already been overcome through the development and characterization of advanced demonstrators. They pave the way towards the "domain specific chiplet on active interposer" route for 2030 as predicted by the International Roadmap for Devices and Semiconductors.

With chiplet-based design, the business model moves from a soft IP business to a physical IP business, one in which physical IP with new interfaces is delivered to a new actor who integrates it with other outsourced chiplets, tests the integration and sells the resulting system. **According to Gartner, the chiplet market size (including the edge) will grow to $45B in 2025** and supporting chiplet-based design tools are available. A challenge is that the chiplet eco-system has not yet arrived due to a lack of interoperability between chiplets making chiplet reuse difficult. For instance, it is not possible to mix an AMD chiplet with a XILINX one to build a reconfigurable multi-core SoC. **Die-to-Die (D2D) communication is the "missing link" to leverage the chiplet-based design ecosystem, and its development in open source would enable a wide usage and could become a "de-facto" standard.**

The Die-to-Die interface targets a high-bandwidth, low-latency, low-energy ultra-short-reach link between two dies. Different types of interfaces exist and the final choice for a system lies in the desire to optimize six often competing, but interrelated factors:

1. Cost of packaging solutions

2. Die area per unit bandwidth (square mm per Gigabits per second)

3. Power per bit

4. Scalability of bandwidth

5. Complexity of integration and use at the system level

6. Realizability in any semiconductor process node

The ideal solution is an interconnect technology that is infinitely scalable (at fine-grained resolution), low power, area-efficient, totally transparent to the programming model, and buildable in a low-cost silicon and packaging technology. There are two classes of technologies that service this space:

- Parallel interfaces: High-Bandwidth Interface (HBI), Advanced Interface Bus (AIB) and "Bunch of Wires" (BoW) interfaces. Parallel interfaces offer low power, low latency and high bandwidth, but at the cost of requiring many wires between Die. The wiring requirement can only be met using Silicon interposer or bridging technology.
- Serial Interfaces: Ultra Short and eXtra Short Reach SerDes. Serial interfaces significantly reduce the total number of IOs required to communicate between semiconductor chips. They allow the organic substrate to provide the interconnection between dies and enable the use of mature System-in-Package technology.

One difficulty of the Die-to-Die approach is that no communication standard currently exists to ensure interoperability. In early 2020, the American Open Compute Project (OCP) initiative addressed the Die-to-Die standardization by launching the Open Domain Specific Architecture (ODSA) project that aims to bring more than 150 companies to collaborate on the definition of different communication standards suitable for inter-die communication.

Recommendation - The chiplet-based approach is a unique opportunity to leverage European technologies and foundries creating European HW accelerators and an interposer that could leverage European More-than-More technology developments. To achieve this, inter-operability brought by open source HW is key for the success, together with supporting tools for integration, test and validation.

The SoC infrastructure for "chiplet-based" components will require PHY and MAC layers of chiplet-to-chiplet interfaces based on standard and open source approaches. These interfaces could be adapted depending on their use: data for computing or control for security, power management, and configuration. This interposer + chiplet approach will leverage European technologies and even foundries, as the interposer does not require to be in the most advanced technology and could embed parts such as power converters or analog interfaces. The chiplets can use the most appropriate technology for each function (memory, advance processing, support chiplets and interfaces). **Interoperability, that could be brought by open source HW, is key for the success, together with supporting tools for integration, test and validation.**

Regarding the connection of SoCs to external devices, some serial interfaces are quite mature in the microcontroller world and there is little differentiation between vendors in the market. It would make sense to align on standard implementations and a defined set of features that can be used by different parties. Open source standard implementations could contribute to the distribution of standards. However, there are also domain specific adaptions that require special features which would make it hard to manage different implementations.

5.5 Supporting Software

The software infrastructure necessary for a successful hardware ecosystem contains Virtual Prototypes (Instruction Accurate and Clock Accurate simulators), compilers and linkers, debuggers, programmers, integrated development environments, operating systems, software development kits and board support packages, artificial intelligence frameworks and more. Indeed, **the idea of open source originates from the software world and there are already established and futureproof software projects targeting embedded systems and hardware development such as LLVM, GDB, OpenOCD, Eclipse and Zephyr**. The interoperability and exchangeability between the different parts of the SW infrastructure are important fundamentals of the ecosystem.

5.5.1 Virtual Prototypes

Virtual Prototypes (VP) play a major role in different phases of the IP development and require different types of abstraction. They can range from cycle accurate models which are useful for timing and performance estimations, to instruction accurate models, applicable for software development, design-space exploration, and multi-node simulation. **Independent of the VP abstraction level, these platforms should strive towards modelling the IP to be functionally as close to real hardware as possible**, allowing users to test the same code they would put on the final product.

As modeling IP is usually a task less complex than taping out new hardware revisions, VPs can be effectively used for pre-silicon development. Modeling can be either done in an abstract way, or using RTL, or by mixing these two approaches in a co-simulated environment. VPs can bring benefits not only to hardware manufacturers that want to provide software support for their customers but also to customers in that they can reuse the same solutions to develop end products. Having models for corresponding open source IP could be beneficial for establishing such IP. With this in mind it would be reasonable to provide a permissively licensed solution, allowing vendors to close their non-public models.

5.5.2 Compilers and Dynamic Analysis Tools

Compilers significantly influence the performance of applications. Important open source compiler projects are LLVM (Low Level Virtual Machine) and GCC (GNU Compiler Collection). The **LLVM framework is evolving to become the 'de facto' standard for compilers**. It provides a modular architecture and is therefore a future-proof solution compared to the more monolithic GCC.

MLIR (Multi-Level Intermediate Representation) is a novel approach from the LLVM framework to building a reusable and extensible compiler infrastructure. MLIR aims to address software fragmentation, improve compilation for heterogeneous hardware, significantly reduce the cost of building domain specific compilers, and aid in connecting existing compilers together. Such flexibility on the compiler side is key to providing proper software support for the novel heterogeneous architectures made possible by the flexibility and openness provided by RISC-V. Note that both LLVM and GCC include extensions such as AddressSanitizer or ThreadSanitizer that help developers to improve code quality.

In addition, there are various separate tools such as Valgrind[35] and DynamoRIO[36] that strongly depend on the processor's instruction set architecture. Valgrind is an instrumentation framework for building dynamic analysis tools to detect things like memory management and threading bugs. Similarly, DynamoRIO is a runtime code manipulation system that supports code transformations on any part of a program while it executes. Typical use cases include program analysis and understanding, profiling, instrumentation, optimization, translation, etc. **For wide-spread acceptance of RISC-V in embedded systems, it is essential that such tools include support for RISC-V.**

Similar to hardware components, for safety-critical applications compilers must be qualified regarding functional safety standards. Here, the same challenges and requirements exist as for hardware components. For this reason, today mainly commercial compilers are used for safety-critical applications. These compilers are mostly closed source.

5.5.3 Debuggers

In order to efficiently analyze and debug code, debuggers are needed that are interoperable with chosen processor architectures as well as with custom extensions. Furthermore, debuggers should use standard open source interface protocols such as GDB (GNU Project Debugger) so that different targets such as silicon, FPGAs or Virtual Prototypes can be seamlessly connected.

5.5.4 Operating Systems

Chip design consists of many tradeoffs, and these trade-offs are best validated early by exercising the design by system software. For instance, when providing separated execution environments, it is a good idea to validate early that all relevant shared resources are separated efficiently. While this in principle should be easy it can be surprisingly difficult as highlighted by the Spectre/Meltdown vulnerabilities.

5.5.5 Real Time Operating System (RTOS)

Most of the common RTOS have already been ported to RISC-V (**https://github.com/riscvarchive/riscv-software-list#p17_real-time-operating-systems**), including the most popular open source options such as FreeRTOS and Zephyr RTOS. Even Arm's mbed has been ported to one RISC-V platform (**https://github.com/GreenWaves-Technologies/mbed-gapuino-sensorboard**), though mainline/wide support will most likely not happen. One of the key aspects of the OS is application portability. This can be achieved by implementing standard interfaces like POSIX which does not lock the software into a certain OS/Vendor ecosystem. There are many RTOS, but two key examples of relevance are:

Zephyr RTOS - Zephyr has been aligning with RISC-V as a default open source RTOS option. Currently RISC-V International itself is a member of the Zephyr project, along with NXP and open hardware providers Antmicro and SiFive. Zephyr is a modular, scalable RTOS, embracing not only open tooling and libraries, but also an open and transparent style of governance. One of the project's ambitions is to have parts of the system certified as secure (details of the certification process and scheme are not yet established). It is also easy to use, providing its own SDK based on GCC and POSIX support. The RISC-V port of Zephyr covers a range of CPUs and boards, both from the ASIC and FPGA worlds, along with 32 and 64-bit implementations. The port is supported by many entities including Antmicro in Europe.

Tock - **Tock provides support for RISC-V.** Implemented in Rust, it is especially interesting as it is designed with security in mind, providing language-based isolation of processes and modularity. One of the notable build targets of Tock is OpenTitan. Tock relies on an LLVM-based Rust toolchain.

5.5.6 Hypervisor

A hypervisor provides virtual platform(s) to one or more guest environments. These can be bare-bone applications up to full guest operating systems. **Hypervisors can be used to ensure strong separation between different guest environments for mixed criticality**, such platforms also have been called Multiple Independent Level of Safety and Security (MILS) or a separation kernel[37]
. When a guest is a full operating system, then that guest already uses different privilege modes (such as user mode and supervisor mode). The hypervisor either modifies the guest operating system (paravirtualization) or provides full virtualization in "hypervisor" mode. **RISC-V is working on extensions for hypervisor mode, although these are not yet ratified.** Some hypervisors running on RISC-V are listed at **https://github.com/riscvarchive/riscv-software-list#p17_hypervisors-and-related-tools**.

At the moment, a number of hypervisors for critical embedded systems exist, provided by non-European companies such as Data61/General Dynamics, Green Hills, QNX and Wind River, and in Europe by fentiss, Hensoldt, Kernkonzept, Prove & Run, Siemens and SYSGO. Many of these are being ported/or could be ported to RISC-V. A weakness is that these hypervisors usually have to assume hardware correctness. An open RISC-V platform would offer the opportunity to build assurance arguments that cover the entire hardware/software stack. **Any new European RISC-V platform should be accompanied by a strong ecosystem of such hypervisors.**

Recommendation - If publicly funded research in the open hardware domain makes available some core results such as the used software/hardware primitives (such as HDL designs and assembly sequences using them) under permissive or weakly-reciprocal licenses, then these results can be used by both kinds of systems.

In the field of system software such as RTOS/hypervisors, currently there are several products which are closed source and have undergone certifications for safety (e.g., IEC 61508, ISO 26262, DO-178) and security (e.g., Common Criteria), and others which have not undergone certification and are open source. The existence of value chains as closed products on top of an open source ecosystem can be beneficial for the acceptance of the open source ecosystem and is common to many ecosystems. An example is the Linux ecosystem which is used for all kinds of closed source software as well.

5.5.7 NextGen OS for Large Scale Heterogeneous Machines

To address the slowdown of Moore's law, current large-scale machines (e.g., cloud or HPC) aggregate together thousands of clusters of dozen of cores each (scale-out). **The openness of the RISC-V architecture provides multiple grades of heterogeneity, from specialized accelerators to dedicated ISA extensions, and several opportunities for scalability (e.g., large-scale cache coherency, chiplets, interconnects, etc.) to continue this trend (scale-in)**. However, manually managing the hardware-induced heterogeneity of the application software is complex and not maintainable.

Recommendation - There is a need for research on RISC-V flexibility to revisit how to design together hardware and operating systems in order to better hide the heterogeneity of large machines to users, taking into account potential disruptive evolutions (non-volatile memory changing the memory hierarchy, direct object addressing instead of file systems, Data Processing Units to offload data management from processors).

5.5.8 Electronic Design Automation (EDA) Tools

Implementing a modern design flow requires a significant amount of EDA tools as shown in **Figure 5.5** – see[38] which states **"With billions of transistors in a single chip, state-of-the-art EDA tools are indispensable to design competitive modern semiconductors"**.

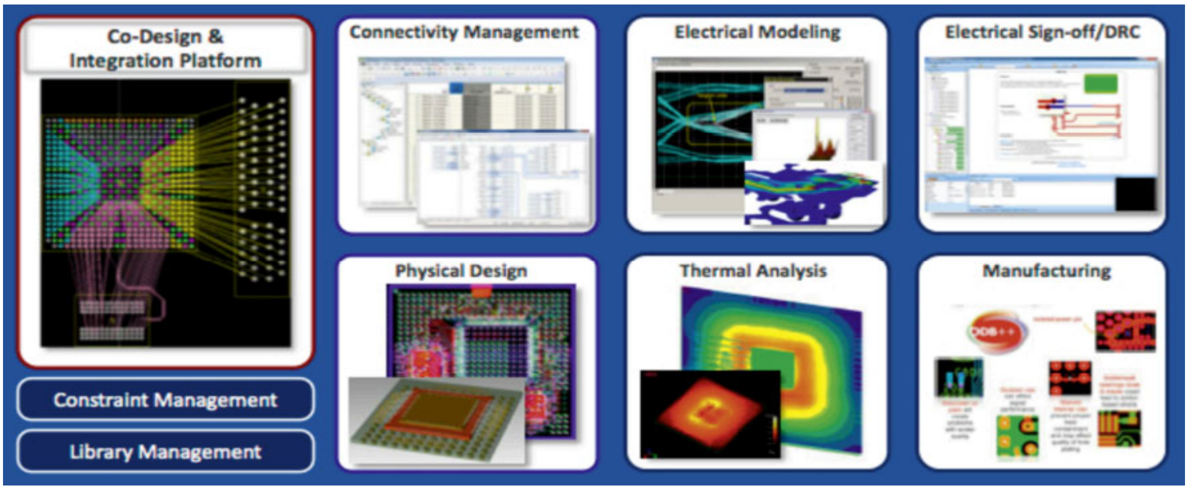


Figure 5.5 - Xpedition Package Integrator Suite (Source: Siemens)

Traditionally, EDA has been dominated by mostly US-based closed source commercial vendors. With projects like Verilator open source activities for specific parts of the design flow have also started to gain traction. However, open source support is still far from allowing a competitive fully open source design flow, especially when targeting digital design in advanced technologies, necessitating the co-existence of the existing commercial tools and upcoming open source ones for many years to come.

Open source tools are essential for introducing new companies and more developers into the field; especially developers with a software background who can bring in innovation in hardware-software co-design. Developers typically do not need to license their daily tools anymore and can freely work together across teams and organizations using massive collaboration hubs such as GitHub or GitLab. These benefits and capabilities need to be enabled via open source tooling for the sector to keep up with the demand for talent and innovation. A vital EDA community already exists in Europe with companies focusing on point solutions within the broader semiconductor flow. **Significant investment into open source tooling as well as cross-region collaboration is needed to energize the sector.** Contrary to common belief, the current EDA giants stand to benefit significantly from open source tooling investment, as there will be continued need for large, experienced players while open source solutions will enable new use cases and provide improvements in existing flows. New business opportunities will be created for the existing EDA players by incorporating new open source development. The recent acquisition of Mentor (one of the three leading EDA companies) by Siemens, means that it joins the European EDA community and can collaborate with the local ecosystem to support European sovereignty. The top EDA companies spend USD $1 Bn+ annually on R&D costs to continue innovating, so to provide meaningful progress in the open source space, continued investment from the public sector and cross-border collaboration are needed to bridge the gap. **Open source EDA should be a long-term goal in order to further the European sovereignty objectives, but existing European proprietary EDA will need to be utilized when necessary in the short to mid-term due to the significant investment that would be required to create a competitive, full flow open source EDA solution.**

Recommendation - Current proprietary EDA tools can help open source design IP development get off the ground if they are available at sufficiently low cost and provided the tool licensing does not restrict the permissive open source usage of the developed design IPs, but a sustained and long term investment into open source tooling is needed to build a sustainable ecosystem.

Large semiconductor and system companies use their pool of proprietary EDA licenses to design, verify and get their open source-based designs ready for manufacturing with the expected productivity and yield. However, they may also be interested to introduce software-driven innovations into parts of the flows and can benefit greatly from the economies of scale of open source, enabling large teams to use the same tools free of charge. Smaller companies and research organizations need access to a comparable level of professional EDA tools which the large EDA vendors will most likely provide through a variety of business models (Cloud offering, specific terms for start-up, research licenses, Software as a Service (SaaS), etc.). The EC is continuing to invest to build a European open source EDA tooling ecosystem, encouraging open interchange formats and making sure current tools do not introduce restrictions on utilization on open source hardware designs independently from the open source hardware license used.

Considering the safety requirements of some of the IPs under discussion here, open source offers a unique possibility to create transparent, auditable processes for ensuring safety. Much like in the case of open source software, common procedures and pooling efforts through oversight bodies (such as the Zephyr RTOS or RISC-V Safety Committees) can be used to provide safety certificates or packages to reduce the burden of safety compliance off the shoulders of developers.

The key to enabling open source tooling in the EDA space (which will most likely also benefit existing, proprietary vendors) is in enabling specific components of the proprietary flows to be replaced by open source alternatives which can introduce point innovations and savings for the users. **This should be encouraged by focusing part of the EU investment on interoperability standards which could allow the mixing of open and closed ASIC tools**, much like the FPGA Interchange Format driven by CHIPS Alliance is doing for FPGAs. RISC-V is another example in the hardware space of how closed and open source building blocks can coexist in the same space and reinforce each other as long as there are common standards to adhere to. EDA tools are exploited in various parts of the development chain including lifecycle management, architecture exploration, design and implementation as well as verification and validation.

5.5.9 Lifecycle Management

A state-of-the-art development process comprising continuous integration and continuous delivery/deployment is at the core of typical projects. A key aspect is requirements traceability, both on the core and up to the system level to ensure that the verification of the requirements can be demonstrated. This is crucial for certification of safety (e.g., ISO 26262) where an IP can be used as a safety element by rigorously providing the requirements and then by ensuring that they are satisfied during integration. Another useful data point for users to decide if they should trust a particular configuration of an IP is the "proven in use" argument. This necessitates collecting data on which configurations of the IP has been taped out in given projects and the collection of related errata from the field. Here open source hardware provides the opportunity to propose better traceability metrics and methods than proprietary counterparts. Complete designs can be shared and even manufacturing information on proper processes and good practices, as well as lifecycle management for open source hardware. This does not preclude the “out of spec” use of open source hardware in other domains such as low critical applications and education.

EDA tooling needs to support tracing the standardized verification metrics from IP level to system level. To successfully span the hierarchies from core to system level, a contract-based design is crucial, allowing to share interface contracts along the supply chain. This needs language and tooling to make it accessible to architects and designers. The open source toolchains, IPs and verification suites allow scaling the continuous integration/continuous deployment systems in server infrastructure reducing the build and test time. Proprietary solutions often require dedicated licensing infrastructure effectively preventing it from being used in scalable, distributed testing infrastructure. Different licensing and/or pricing of proprietary EDA tools for open source development may help to leverage the existing technology at the early stage in the open source development, before the open source alternatives are readily available.

5.5.10 Architecture Exploration

It is crucial to make the right architecture choices for the concrete application requirements before starting down a particular implementation choice. This requires tool support to profile at a high abstraction level of the application before software is developed. Usually, hardware is first modeled as a set of parameterized self-contained abstractions of CPU cores. Additional parameterized hardware components can then be added and software can be modeled in a very abstract level, e.g. a task graph and memory that get mapped to the hardware resources such as processing units. **Simulating executions on the resulting model allows analysis of the required parameters for the hardware components. This can include some first impressions on power, performance, area for given parameter sets, as well as first assessments of safety and security.** The derived parameter sets of the abstract hardware models can then be used to query the existing IP databases with the standardized metrics to identify matching candidates to reach the target systems goals.

The profiling can provide Power Performance Area requirements for instructions and suggest ways of potentially partitioning between accelerators and the relevant ISA instructions can be explored to meet these requirements. Ideally, a differentiated power analysis can already be started at this level, giving a rough split between different components like compute, storage, and communication. This can then be refined the more detail.

A widely used methodology is Model-Based System Engineering (MBSE) which focuses on using modeling from requirements through analysis, design, and verification. For the application of processor cores, this would mean the use of Domain Specific Languages (DSLs) down to the ISA level to capture the hardware/software breakdown. The breakdown would then be profiled before starting implementation of any design or custom instructions on top of a core. The modeling should lend itself for use in High-Level Synthesis (HLS) flows.

Recently, thanks to the popularization of open source hardware (which pushes more software engineers into the hardware industry) as well as a fast-moving landscape of modern AI software which in turn requires new hardware approaches, more software-driven architecture exploration methods are being pursued. For example, the Custom Function Unit (CFU) group within RISC-V International is exploring various acceleration methods and tradeoffs between hardware and software. Google and Antmicro are developing a project called CFU Playground[39] which allows users to explore various co-simulation strategies tightly integrated with the TensorFlow Lite ML framework to prototype new Machine Learning accelerators using Renode and Verilator.

5.5.11 Design & Implementation

Architecture exploration leads to a generic parameterized model. This requires a modeling language that offers sufficient expressiveness and ease of use for wide acceptance. A challenge is that typically engineers engaged in the architecture exploration process do not come from a hardware background and are unfamiliar with SystemC or SystemVerilog, the established hardware languages. Modern programming languages like Python or Scala are thus gaining traction as they are more widely understood. An architecture exploration process requires the availability of models and an easy way to build a virtual platform and exchange models by implementations as soon as they become available. This procedure enables early HW/SW Co-Design in a seamless and consistent manner.

A traditional Register-Transfer Level development process manually derives RTL from a specification document. To address the large scope of parameterized IPs targeted in this initiative, more automation is required. **The parameterized Instruction Set Architecture (ISA) models for processors lend themselves for a high-level synthesis (HLS) flow where detailed pipeline expertise is not required by the users unless they require the highest performance. This is crucial for enabling a wider audience to design and verify processors.** The design process thus lends itself to a high degree of automation, namely offering some form of HLS starting from the parameterized models. For processors, for example, synthesizing the instruction behavior from an ISA model and thus generating the RTL is feasible for a certain class of processors. The design process should also generate system documentation in terms of:

- System memory map
- Register files of each component
- Top level block diagram
- Interconnection description

The documentation needs to be created in both user and machine-readable forms allowing manual and automated verification.

Recommendation - There is a need for a high degree of automation for adding various monitors in the design flow without having to manually pick a suitable monitor IP or even to design, configure and wire it up from scratch, which can be very error prone.

RTL development flows should rely on continuous integration processes, such as automated checking via linting and Design Rule Checks (DRC). There is a possibility for RISC-V specific linting or DRC. Similar checks can be applied down the flow after synthesis to catch simple mistakes which often result in huge consequences.

It is important to add some monitoring components in the design phase, among other things to address threats that cannot be fully verified during development or that arise because of hardware/software interaction. In addition to an early security threat warning, monitoring is also needed for the safety. Design integrity and performance should also be monitored. While RISC-V offers a generic performance counter mechanism, not much has been standardized, creating an overhead to get specific counters like cache statistics integrated into the toolchain. Common C++ libraries for accessing the counters should be available. Debug logic with a more active role is needed to replace today’s mostly passive monitors. Heterogeneous cores, each with their individual debug features, creates additional software challenges. The generated monitoring and debug data can be evaluated with in-system software or off-loaded from the chip. **In order to evaluate the comprehensive monitor and debug data from such a SoC new analytics and visualization tools will need to be developed. There is also a need for system level and context specific debug tailored to applications like automotive, 5G, HPC.**

5.5.12 Verification & Validation

Verification management needs to be tightly integrated with lifecycle management for traceability from requirements to verification. **Verification should benefit as much as possible from the models produced by architecture exploration creating golden models for verification with coverage goals.** At the system level there are hard challenges like cache coherency and there is a need for portability of tests across different levels. State-of-the art verification approaches use simulation based and formal approaches. The introduction of automated formal verification has made the approach available to typical design and verification engineers allowing them to set up the targeted checks of the application much faster than they would in a simulation-based flow. **The approaches are complementary, lowering the overall verification effort with formal verification applications, while at the same time increasing verification quality in the crucial areas with formal proofs compared to an incomplete simulation.** To further enable a broad user base for custom processors, the verification side of the flow also needs a high degree of automation, for example deriving large parts of the verification from the parameterized models also used on the HLS side to create the implementations. For safety related projects, state-of-the-art tools and methodologies must be used driving the use of formal verification.

Cores and IPs should come with a reference flow similar to what is provided by Arm, allowing to re-run the provided verification in various tools, be it proprietary or open source. The CHIPS Alliance is developing an open source RISC-V core code generator/verification framework called RISC-V DV[40] which the OpenHW Group [41] is also using as the base of a simulation-based environment for verification of their cores. The CHIPS Alliance is working towards a fully open source Universal Verification Methodology verification flow based on Verilator [42]. However, the OpenHW environment still needs a proprietary simulator to run the full verification. Only a small subset of the verification suite can be run on the open source EDA tool. **For open source design IPs with industrial strength verification, there is no short or mid-term availability of an open source EDA tool suite that provides simulation-based and formal verification. Achieving an industrial strength verification with open source EDA verification tools is a long term goal.**

Another example of open source verification work can be found in the OpenTitan project[43] which provides a valuable example of an open source, continuous delivery system whose coverage and status can be traced for every commit. Software-based frameworks like cocotb are also gaining traction, especially with engineers and teams with a software background, and while they are incompatible with traditional Universal Verification Methodology (UVM) style verification, many new open source IP implementations adopt them. Hardware description languages based on modern programming languages like migen/nMigen (Python based), CHISEL, SpinalHDL (Scala based) provide their own simulation and verification flows. Since the above languages are derived from modern programming languages, they can easily reuse testing methodologies known from the software world.

Recommendation - For open source EDA tools, establishing collaboration models realizing a sophisticated design and verification process is an important topic. Maintaining an EDA tool for safety-critical applications after an initial release also requires a significant amount of both manpower and computing resources to ensure persistent tool quality and up to date safety collaterals.

As an example Antmicro's Renode while used mainly for software development and testing, provides a means to create complex simulation environments based on the Hardware Description Language (HDL) code of both cores and peripheral IPs. This creates an easy way to test IP in complex software scenarios, instead of synthetic, hand-crafted tests. In addition, it gives the possibility to work on software development in the pre-silicon phase of an ASIC project reducing the overall time-to-market.

Recommendation - For the standard protocols used to connect the design IPs, the use of Verification IP (VIP) is strongly encouraged. VIP allows all protocol rules on the controller or peripheral sides to be checked, thus ensuring IP blocks from different vendors can properly communicate. Formal VIP is another typical application where formal verification is easy to setup and gives great verification results. A focus should be on the protocols used in the design IPs – be it the AMBA protocols or open source ones like OBI, OCP, or TileLink.

5.5.13 Tool Qualification for Safety-Critical Applications

In the automotive industry, a classification according to ISO 26262 is necessary for all design and verification tools that should be used in a safety-critical environment. If the classification shows that the tool could introduce errors into the design and these errors would not be detected by another step in the design flow, a tool qualification is needed.

This leads to a significant challenge for the quality of the tools. By having a sophisticated development process handling requirements traceability, change management and comprehensive documentation, the work on the user side to qualify a tool can be significantly reduced. Furthermore, proof of comprehensive test suites and therefore the complete verification of the tool is necessary for a tool verification.

5.6 Cross Cutting Requirements

Instruction sets have traditionally been managed by a single "owner" (Intel, Arm) and their evolutions over time have been slow, lagging application requirements. This is because the ISA-owners are not strongly motivated to modify or overhaul their ISA and they hesitate to invest the required R&D and engineering effort, typically engaging only under extremely heavy customer requests. This situation has completely changed with the advent of the RISC-V open ISA, mostly for two reasons:

(I) from a technical viewpoint the RISC-V ISA is designed to be modular and extensible, with adequate provisions for ensuring backward compatibility issues

(II) R&D and commercial efforts to extend the RISC-V ISA for specific application domains can be initiated as community efforts, with cost and risk sharing and can be also used to provide differentiated value-added solutions.

As a consequence of this paradigm shift toward open ISAs enabled by RISC-V, major innovation opportunities are enabled on open cores with enhanced instruction sets. The faster innovation cycle which is now possible by coupling ISA and core enhancements in an open source setting, has been demonstrated in several domains. Notable examples include: open extensions for supporting quantized computations in machine learning, with particular emphasis on deep neural network inference, general digital signal processing extensions (e.g. Single Instruction Multiple Data (SIMD), fixed-precision arithmetic), and security extensions (e.g. Galois field operations, bitwise operators). It is also important to note that extensions that have been successfully prototyped as non-standard ISA enhancements and have gathered wide adoptions, can then be moved toward new standardized parts of the ISA.

Scalability is also another important opportunity created by the RISC-V instruction set architecture. RISC-V does not prescribe a limit to instruction encoding size. In addition to the already standardized 16-bit, 32-bit and 64-bit instructions, it is also possible to specify intermediate (e.g., 48-bit) and even larger instructions sizes: 128-bit instructions are already being worked on in the RISC-V standardization committees.

5.6.1 Safety Certification – Open Standards; Safety-Ready Toolchains

RISC-V is promising for applications in the high-assurance market due to potential cost reductions from easier access to innovation, flexible and rapid design processes, stability and modularity, and availability as white box[44]. When considering safety there is a need to consider both the hardware component level and also the system level taking into account the interplay between hardware and software. At the hardware level the safety-critical hardware components should ideally be developed to sectoral safety standards. For example, in the automotive sector the key standard is the ISO 26262 standard which contains development process requirements in order to avoid systematic and random faults. These address management processes during the development lifecycle, role definitions, hazard analysis, risk assessment and development processes. These include conditions for requirements tracing, confidence levels of tools used for development and verification and verification requirements. Certifying a system that contains hardware components that are not developed according to the ISO 26262 is still possible but elaborate.

Recommendation - In order to make the use of open source hardware components possible in automotive applications, artifacts and methods have to be defined that help with certifying these components. A start could be the implementation of quality-management systems in open source projects.

In addition to the open source availability of source code, the open source availability of verification artefacts (verification plans, test benches, reference simulators, assertions, test sequences...) can be enablers of white-box analyses and certification processes. Notably the OpenHW Group already publishes these verification artefacts[45], however, available documentation of the verification process itself still might not be sufficient especially for the highest assurance levels (ISO 26262 ASIL D, DO254 DAL A...) where additional, and often very costly, verification practices might be needed.

EDA tools can be used for implementing and assessing the safety of hardware components. Automation plays a key role in order to avoid the need for RTL coding and development of verification strategies in the otherwise largely automated process. Similarly, tool support for adding and validating security mechanisms is needed. It is important to be able to assess the impact of safety or security mechanisms as much as possible based on models used as input to the design phase, i.e., before the design is completely done. In addition to assessing single safety/security mechanisms, the assessment at the system level is also crucial, including the fulfilling of critical timing deadlines in real-time safe systems. A key challenge is the huge parameter space of typical IPs as validation is needed for each and every concrete parametrization that is used. **For safety in particular, support for specific certification flows like Failure Modes, Effect and Diagnostic Analysis (FMEDA), is of huge value for integrators.**

At the system level a key issue is to guarantee Worse Case Execution Time (WCET) deadlines for the execution of critical tasks. For any time-critical system, requirements specify maximum response times. An issue is that most hardware manuals now do no not publish steps/cycles/execution times making a-priori determination of execution time difficult. In the past this information was available, e.g., Intel 486/Pentium manuals[46]. However, newer processor (Intel, PowerPC, Arm) manuals do not provide such numbers for multiprocessors like the P4080.[47] A workaround is to replace clock cycle analyses by own empirical time measurements which suffices for average case behavior but is inefficient and/or unsafe for assessing worst-case execution time.

5.6.2 Security

Common Criteria security certifications for simple hardware IP such as smart cards[48] exist, however, for more complex processors, security is still in its infancy and it is not possible to buy a Common Criteria certified general purpose multicore processor. To get around this currently companies have to make liability limiting statements such as "The underlying hardware [...] is working correctly and has no undocumented or unintended security critical side effect on the functions of ..." However, relying on the hardware's documented interface alone can be insufficient, as there may be parts of the hardware-software system architecture that cannot be inspected by the software developer. For instance, **the Spectre/Meltdown vulnerabilities which appeared in 2018 were unexpected for almost all OS vendors.**

In an exhaustive search of hardware-software systems (3800 publicly available security targets and evaluation reports available from **commoncriteriaportal.org**) it is noted that the focus is on describing hardware at a high level without going into the micro-architectural side effects. Often unjustified assumptions on hardware are made and no references were found to key issues such as "branch predict*", "translation lookaside buffer", or "branch history buffer" with very few mentions of "cache flush". Evaluation reports do, however, exist on the analysis of the security properties for public-key cryptography including both VHDL and software analysis. A recent effort on the security side is Accellera's SA-EDI (Security Annotation for Electronic Design Integration) that establishes a link to CWE (Common Weakness Enumeration). **A continuing challenge is long-term security as there is a need to continually innovate to guard against future new attacks.**

5.6.3 The Way Forward - Hardware-Software Contracts for Safety and Security

Safety and security properties need to be known and can be summarized in the concept of hardware-software codesign with contracts as shown in **Figure 5.6**. **In this open hardware allows development of hardware-aware system software and system software-aware hardware.**

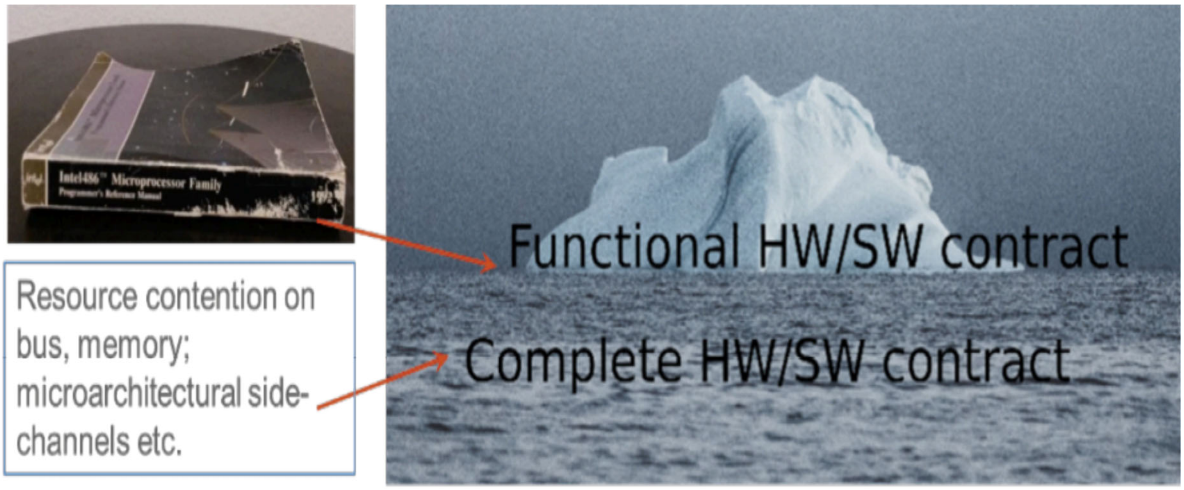


Figure 5.6 - Hardware/Software Contracts

In this approach rely-guarantee relations are made between hardware and software called "hardware-software contracts" [49]. Here there is a key need to bring together hardware providers and software providers to develop and produce a general-purpose open hardware platform or at least core components with access to the full hardware-software contract, allowing safety and security certification. This should be based on an instruction set architecture used in an existing or upcoming ecosystem, such as RISC-V. At the same time a verification methodology that provides reasonable assurance for the platform/its core components is needed. These verification artefacts need to be published to allow them to be used in use cases.

Recommendation - For supporting safety and security there is a need to demonstrate and share verification environments and artifacts of open source hardware at medium assurance levels and also at higher Safety Integrity Levels (SIL), e.g. safety SIL 4/ASIL D, security Common Criteria EAL 6 and higher, firstly addressing simple systems and then moving to complex systems.

5.7 Identified Gaps And Future Needs

The European economy consists of multiple industrial application domains which will be impacted by RISC-V and Open Source in different ways. These application domains include automotive, industrial manufacturing, consumer electronics, telecommunications, health, home applications, aerospace and defense. Within these domains there are some common requirements, and also specific challenges, such as security, safety, privacy and support for artificial intelligence. **Although these application domains can make use of generic common processor families, in many cases domain-specific processor families or specific features are needed to meet performance, reliability, cost, and energy consumption requirements.** In this section a number of common, as well as application/domain-specific gaps and needs are identified through the process chain as shown in **Figure 5.7**.

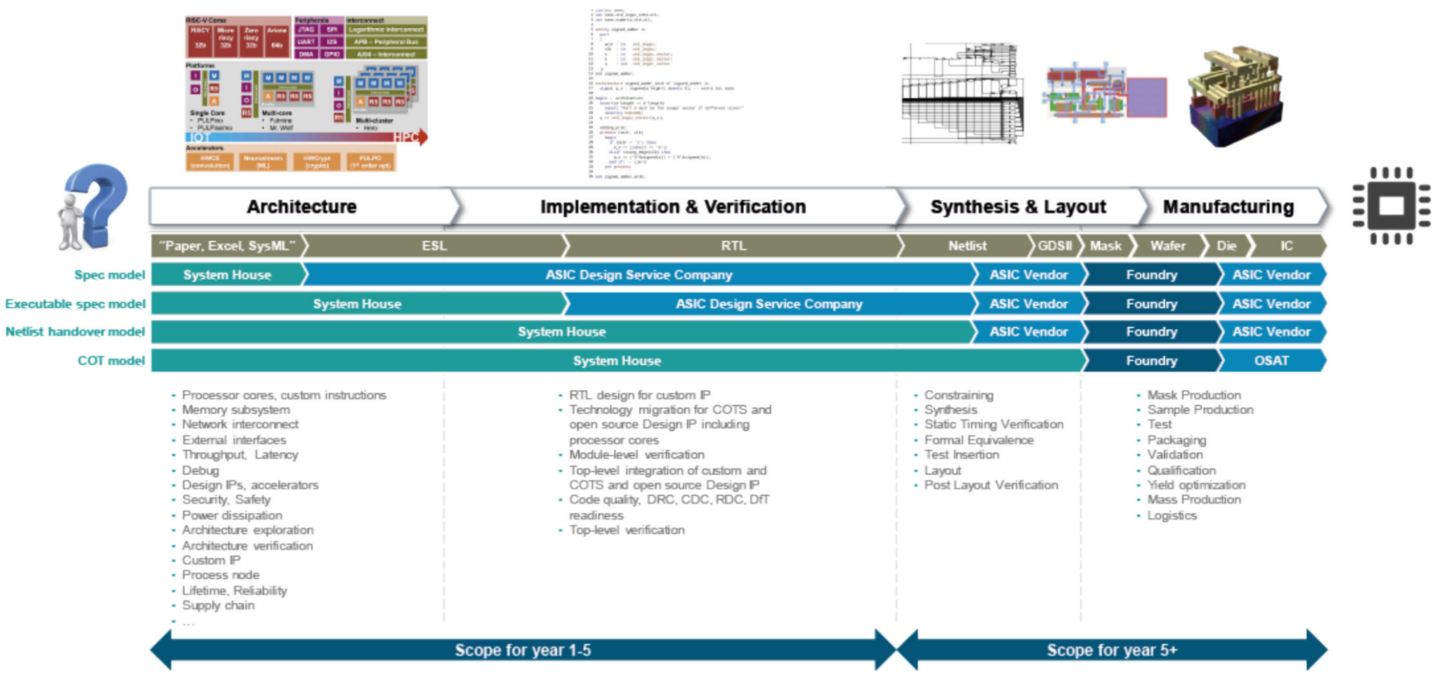


Figure 5.7 - From Architecture to Manufacturing

5.7.1 Processor Design

On the processor side, **European technology already is well established for low-end processors**, which could be considered micro-controller-class cores. Far fewer efforts deal with mid-end cores, which are sometimes called application cores. **The lack of European sovereignty is most pronounced, though, at the high-end, which includes server-grade and HPC cores**. In this area, there is a complete absence of European open source offerings.

5.7.2 Core Microarchitecture

To close these gaps, there is a need to strengthen efforts for both the mid-end and the high-end. In the first case it will be possible to build/improve upon existing open source offerings. For high-end processors, work will have to start mostly from scratch, as only very limited prior work is available even internationally. For instance, the BOOM processor initially developed at UC Berkeley, is still undergoing heavy revisions (e.g., from v2 to v3) and does not appear to be a stable base for future development yet. For the open source Alibaba Xuantie 910 core key documentation is only available in Chinese. A European effort should consider the results of these prior works but will most likely be starting (almost) from scratch for a European implementation. This is especially true when the high-end core should be evolvable towards advanced features such as multi-threading, smart prefetching, multi/many cores. Furthermore, the verification of the complex out-of-order superscalar microarchitectures required for high-end processors is considerably more difficult than the simpler in-order ones in the low-end and mid-end cores. Since Europe has a strong research background in formal verification techniques, the design of a high-end core could be a key opportunity for applying that expertise to practice. For instance, a result of formal verification techniques could be proofs of the absence of side channels.

5.7.3 Hardware Peripherals and Interfaces

In most cases, the processor cores described in the previous section will not operate completely stand-alone but will be integrated in a more complex system having additional components both on-chip and off-chip. Interaction between some of the components requires complex and challenging interfaces that are so-far mainly available as proprietary solutions from commercial vendors (such as physical interfaces to high-speed links or memories). More complex use-cases include the communication between core(s), between cores and accelerator(s), and also to memories (unified memory approach), especially when coherency between distributed local caches has to be covered as well. This problem becomes even more complicated when also considering safety-critical or mixed-criticality aspects in the solution. **Depending on the communication scenario, e.g., number of cores and accelerators, throughput and latency requirements, it may become necessary to upgrade from mostly passive interconnects (matrices of multiplexer-based switches) to active structures such as networks-on-chip (NoCs), which can scale better with more communication partners.** Open source baselines already exist for some of these requirements, e.g., cache coherency between a relatively small number of nodes. However, these do not scale to the more complex accelerator-intensive systems becoming more common today, and do not address the safety-critical or mixed-criticality aspects at all. There is a need for a flexible open source infrastructure for these on-chip communications, which should be easily customizable for the specific functional and non-functional requirements of a concrete SoC. In addition to the physical IPs there is also a need for tools to help in the design and the dimensioning of the infrastructure, the simulation, validation and verification views.

The complexity of off-chip interfaces spans an extremely wide spectrum, mostly depending upon the desired bandwidth. For low-speed interfaces, such as UART, SPI, I2C, I2S, etc., a wide variety of open source blocks exist. However, for higher speed (and more complex) interfaces such as USB3, 1G Ethernet, (LP)DDR2+ dynamic memories, and PCI Express, the design challenge increasingly moves from the digital into the analog domain. These interfaces require not only increasingly complex digital logic controllers (memory, PCI Express), but also high-speed off-chip communications using the appropriate I/O logic and so-called fast PHY/SerDes interfaces to achieve the required signaling rates. In general, the design of the PHY/SerDes blocks requires knowledge of low-level electrical characteristics of the underlying chip fabrication technology that is often only available under strict Non-Disclosure Agreements (NDA).

Many relevant system-on-chips that go beyond simple microcontrollers require these high-speed interfaces, e.g., to work with AI/ML datasets exceeding the (at most) single-digit megabytes of static memory storage that can economically be supported on-chip, or to communicate with other peripherals attached via PCI Express (e.g., to mass storage), or with other systems over the network. The lack of open source blocks for the tasks makes it extremely challenging for innovators to prototype solutions, e.g., to attract further investment, as **a commercial license for a DRAM interface (controller and PHYs, commonly provided by an extremely small number of non-European companies) may well cost over half a million USD.**

Recommendation – Effort should focus on areas where there is a current lack of technologies and tools to design and implement modern system-on-chips in order to strengthen European Digital Sovereignty.

To strengthen European Digital Sovereignty, it is essential that this scarcity of key technologies essential for the design and implementation of modern system-on-chips is alleviated. However, due to the confidentiality issue discussed above, realizing these high-speed mixed-signal peripheral interfaces blocks under an open source model is more complex than for the purely digital blocks (e.g., processor cores, many accelerators). **A potential approach could be to design the interface as much as possible in a portable, fabrication-process independent manner.** Only for those parts where it is unavoidable would a concrete ASIC process design be used, ideally one with fabrication facilities within Europe. The process-independent parts could then be customizable and freely open sourced under any suitable license. The process-specific parts, though, will require another approach, e.g., a multi-way NDA between the potential user, the foundry, and the designer of the interface IP. For any project involving such a mix of open and proprietary technologies, concrete work packages and tasks should be allocated to investigate and/or propose licensing strategies/legal frameworks for dealing with this scenario.

In addition to the benefits of making the functionality of high-speed interfaces available to innovators in an open source manner, there is also a secondary benefit that due to their open nature, it is possible to analyze and verify **open source peripherals and interfaces in far greater**

depth than would be possible with closed-source offerings. This is required as the communication protocol is often quite complex making it difficult to ensure that the IP is fully compliant with specifications or standards. This can also enable advances in the areas of security and safety-critical systems, e.g., where interference effects such as bus contention, or side-channels between applications, need to be closely monitored. This especially applies to architectures mixing different criticality levels on the same hardware. For instance, using the design techniques described above (highly focused on customizability), it should be possible to parametrize an interconnect IP block so that selected communications between specific execution environments are guaranteed to be performed in an interference-free manner.

5.7.4 Non-Functional Requirements

In addition to the design/improvement of the actual cores there is also the need to fulfill two key non-functional requirements: **As many of the European users of these cores come from industry (e.g., automotive, Industry 4.0, etc.) where** safety/security/reliability **and the associated certifications are often more important than the performance or power efficiency that would be focus of purely data-center oriented cores. Second, due to the domain-specific needs of these industries, there is a requirement to easily** customize the operation **of the core for applications, e.g., by adding specialized instructions or compute accelerators for operations in AI/ML, cryptography, signal and image processing, custom arithmetic, communications and storage, etc.** It is this need for flexibility with regard to safety/security/reliability and extensibility with custom operations that will be a key differentiator between European core design efforts and the more focused ones (e.g., for data centers, mobile communications, etc.) from existing players. Ideally, by employing state-of-the-art design methodologies leveraging, e.g., new hardware construction languages such as Chisel, Spinal, Bluespec, it should be possible to design the new cores as highly parametrizable base designs to/from which new features can easily be added/omitted as required for an industrial use-case. Note that these hardware construction languages are all open source and based on underlying languages such as Scala and Haskell, which historically have a strong European involvement. However, to fully unlock their productivity gains for industrial-strength use, verification support for these new languages needs to be significantly improved.

5.7.5 Software tools

On the software side, the existing baselines are much more mature than on the hardware side. The huge open source software ecosystem, grown around efforts such as GCC, LLVM, and Linux, has advanced the state-of-the-art considerably over the last 30 years. However, to fully exploit the open RISC-V ISA and the capabilities of the customizable processor cores described above, there are two areas where there is a need for key improvements. The first one is to support the customizability of the new cores also in the associated software. For instance, automatically integrating custom instructions and accelerators into the compilers and operating systems/middleware/hardware abstraction layers. **Without this kind of support, the capabilities of the newly designed hardware will remain inaccessible to most software developers**. The second area plays again to the traditional European strength in safety/security-critical systems. Again, to exploit safety/security-hardened cores the related software must also be extended, e.g., to be suitable for certification under standards such as DO254/DO178 or ISO 26262. This effort can also profit from new technologies, e.g., the combination of processor trace capabilities with dynamic verification techniques that, together, can monitor/enforce safety and security constraints in a running system.

5.7.6 Electronic Design Automation Tools

Electronic Design Automation tools are required to create new hardware. **Central to any hardware design targeting an integrated circuit are the ASIC front-end and back-end implementation tools.** At the front-end, these tools take a description of the function or structure of the electronic circuit, for digital logic circuits such as the RISC-V cores (most commonly in hardware design languages such as Verilog and VHDL), or the more modern hardware construction languages and translate it into more basic building blocks (e.g., logic gates and state elements) by logic synthesis. In the next step, these basic building blocks are mapped to the available hardware blocks in a given chip fabrication process (e.g., standard cells from a library). The geometrical arrangement of these cells is then determined by placing them on a 2-D plane (or in the future: a 3-D volume) and computing the best way to establish the required connections using a routing algorithm. Today, most of this software pipeline relies on proprietary tools developed and sold by a very small number of vendors, mostly from the US. **To strengthen European digital sovereignty in this area, it is crucial that open source solutions are created as an alternative.** Fortunately, initial open source efforts such as the OpenROAD project have already sprung up, sometimes utilizing European-developed tools for core functionality, such as the logic synthesis and mapping steps. These existing efforts could be used to bootstrap advances that would increase robustness of the tools, improve the quality of results by the integration of newer algorithms (e.g., AI/ML-based), and allow fundamental innovation on better tool support for the topics of specific interest to the European microelectronics industry, such as safety/security/reliability. **Additionally, the** practical usability **of such an open source ASIC implementation tool flow should be strengthened, e.g., by ensuring its full interoperability with the process technologies supported at chip manufacturing facilities located in Europe.**

5.7.7 Supplemental System-Level Tools

With modern systems-on-chip becoming ever more complex, there is a growing need for additional steps in the EDA tools workflow supplementing the ASIC implementation flow. These include, e.g., the ability to more abstractly describe the functions of a digital circuit at the purely behavioral level, and then employ a process called high-level synthesis to compile this behavioral description down to the traditional descriptions discussed above. Open source tools for high-level synthesis already exist, however, the approach has proven useful so far mainly for very specific use-cases, such as quickly creating specialized accelerators, but not so much for creating high-quality general-purpose processor cores. Thus, a greater focus should be applied to tools enabling more immediate practical benefits to a larger group of users. Two examples include architecture exploration/optimization as well as automated system-on-chip composition.

Architecture Exploration – **Architecture exploration takes place prior to chip design to determine the best hardware architecture for achieving the design goals**. It employs a number of techniques, such as virtual platforms and software frameworks to model application workloads, to quickly iterate on different hardware architecture choices before committing to a single one which is taken forward to the laborious and costly ASIC implementation and fabrication process. It is also required in the case of a parametrizable design to determine which parameters are best suited for the final design. Despite the importance of this step, as the architectural choices made for an actual chip implementation can have far-reaching consequences for the success of the entire endeavor, only very limited tool support, mainly from a small number of vendors of proprietary software, is currently available. Thus, it would be highly beneficial if a flexible tool framework for architecture exploration could be created in an open source fashion. The open source nature of the framework would then allow innovators to focus on key areas, e.g., the use of AI/ML techniques to optimize entire systems and help the designer to define the best system architecture, as well as alleviate the need to expend R&D effort on laborious engineering tasks, such as visualization, simulator interfaces, etc.

Automated System-on-Chip Composition – There is also a need for support for automated system-on-chip composition. Modern SoCs are, to a large degree, created by composing building blocks (also called IP blocks), where a block may be, e.g., a complete processor core, cache level, or memory controller. Although standards exist for the machine-readable description of individual blocks (e.g., IP-XACT), composing them into an entire system-on-chip encompassing the required hardware/software interfaces, is still mainly performed manually, at best aided by GUIs or low-level generic scripting languages such as Tcl. There is a need for more abstract descriptions of complete SoCs, above the view offered by formats such as IP-XACT. **A new class of EDA tool is needed that can interpret these more abstract descriptions to allow complete system-on-chips to be composed, including optimized design choices, e.g., for interconnect and bus protocol optimization.** Such an SoC composition tool could also be integrated with the architecture exploration flow discussed in the prior paragraph as one step of an automated design-space exploration mechanism. As before, having the SoC composition tool available as open source would enable both more innovation in the operation of the tool itself, but also allow its easier integration with other automation tools, such as architecture exploration tools. In this way, innovators in the hardware space could concentrate on their specific unique contributions (e.g., new AI/ML or cryptography accelerators), and expend less valuable engineering resources (especially for startups) on non-innovative, but complex engineering tasks such as manual SoC composition.

5.7.8 Verification

An overarching topic across all of the different hardware design and implementation steps already discussed is verification. This occurs at all levels, e.g., from architecture specifications, such as the formulation of custom instructions of a processor, down to the actual physical-level silicon design. **One of the crucial aspects for the success of design IPs is the completeness of specifications, validation of the architecture and verification quality of the resulting IP.** Efficient design exploration tools and simulators allow to identify and specify the best architecture parameters for an IP (some tools even allow to generate VHDL or Verilog code for particular specific domains). The earliest that bugs or deviations from the specifications are identified, the less expensive they are to correct. State of the art industrial strength functional verification is required for ASIC tape-outs, even more so in safety critical applications mandated by standards like ISO 26262. Such industrial strength functional verification cannot be achieved by constrained random simulation alone but requires formal verification and emulation in the tool portfolio. A tool agnostic verification flow spanning these three elements is the basis for achieving industrial strength verification. This flow should be compliant with safety standards to allow using the results for safety certification. Such a complete verification includes a safety certification-ready verification plan, shareable/reusable requirements and artifacts, requirements tracing through to the verification plan and the tests verifying them.

Verification of the embedded software stacks using virtual platforms and hybrid platforms (mixing high-level models, e.g., SystemC and RTL) in the system context, if it comes after RTL design & verification is too late. **Tools enabling this task to be performed much earlier in the design cycle are thus essential to cut design & verification time and cost**.

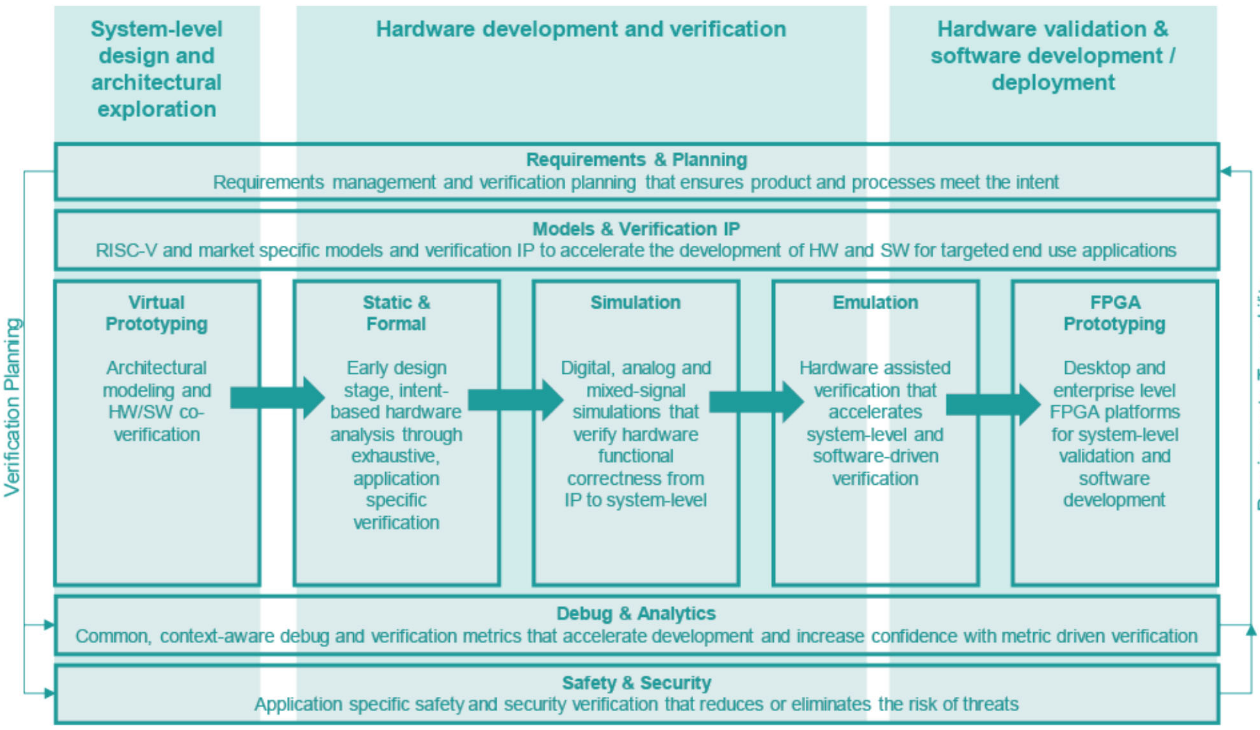


Figure 5.8 - V Model at System and Hardware Level

State of the art IP design also uses tools for linting or design rule check (DRC), clock domain crossing (CDC), reset domain crossing (RDC), X propagation and formal auto-checks. Additional tools may also be mandated for achieving compliance with downstream flows. **In order to achieve highest RTL verification quality in a manageable amount of time, automation is key.** The verification flow needs to be as automated as possible, from continuous integration (CI) to, for example, formal verification apps targeted at certain aspects of the design IP. **In the context of design IPs around RISC-V cores and other subsystem components, an automated formal verification is needed for those parts that are complex and hard to verify.** At the core, there is a need to verify a parameterized RISC-V core with custom instructions against its expected ISA. This requires checks that all instructions and registers behave according to the ISA and no other instructions or registers exist that could offer backdoors for malicious attackers. Beyond the mere core, other areas of interest for an automated formal app are caches, interconnect, prefetch, or consistency for memory systems with several hardware threads or cores.

In addition to functional verification, security topics are starting to play a more and more crucial role. Unexpected discovery of attacks like Spectre or Meltdown on state-of-the-art CPUs highlighted a verification gap that calls for a formal approach that can actually find all kinds of such attacks instead of patching designs after the fact whenever a new attack is found[50]. Design IPs have additional security requirements that must be part of the verification flow, including traceability. Verification goals may also apply for security, resulting in a set of security checks applicable to an IP. This is again a domain where automated formal apps can offer great benefits. In terms of standards, the field of security is less mature than safety. **Nevertheless, the verification flow should also allow integration with standards like Accellera's Security Annotation for Electronic Design Integration (SA-EDI) and offer the corresponding automation.**

On top of the RTL, there is also system software that has to interact correctly with the design in order for the system to work properly with respect to safety and security. **Here, there is a need to focus on boot code and firmware.** Both are critical components with large security impact, and both are very close to the hardware. Again, formal based apps could be used to investigate these software parts in conjunction with the hardware. Formal apps can identify unexpected sources that can influence the boot process or spot parts of firmware code violating the programming rules.

5.8 Identified Gaps Summary

Conception Stage

- There is a need for tools allowing fast simulation and exploration of the design space configuration, and automated system-on-chip composition. Commercial tools exist already as well as open source ones, but they can be improved using new approaches such as ML/AI.
- The community would benefit from a repository of open source models and new technologies that can be used in both open source and commercial tools. New technologies, for example using Artificial Intelligence, can be specifically developed to increase the productivity of architects and designers.
- Simulation models for architecture exploration and optimization should be made available.
- Tools to support early verification of embedded software are needed.
- An open source repository of peripheral models, interconnect, etc., with different abstraction levels is needed.

Automated Composition/Optimization Tools

- Tools are needed to support complex heterogeneous SoCs composed of a mix of RISC-V cores, accelerators and arbitrary design IP to increase productivity in SoC design.
- Integration with existing frameworks and development of extended capabilities including the generation of HW hardware/software interfaces and low-level FW is required.

Functional Verification

- There is a need for an automated industry strength verification, including formal apps, of caches, interconnect, and memory consistency in multi core systems.
- An automated tool-agnostic verification flow integrating simulation-based, formal and emulation-based verification compliant with safety standards is essential to lower the adoption barrier of open source design IP and productivity of SoC design.
- Point technologies need to be integrated into automated flows to hide complexity of verification away from users.
- An open source verification database including artifacts along lifecycle of design and verification of IPs for AI/ML-enabled EDA tool R&D is required.
- To ensure system software in sync with the RTL, additional automated formal verification apps for boot code and firmware together with the RTL are crucial.
- Design IP code quality is crucial on top of functional correctness. Integration of Lint/DRC, CDC, RDC, formal auto-checks, compliance with downstream flows, CI/CD flows into open source design flow management frameworks and point technologies for individual analysis steps is needed.

Functional Safety and Security

- Reference flows should be defined that describe and demonstrate the state-of-the-art methodologies in order to deliver products that maximize confidence and optimize development efficiency.
- Automated tools are needed to provide quantifiable verification and validation of safety and security allowing developers to focus on their core differentiation technologies.
- Tools, methodologies and associated work products should be developed to that they enable sharing of information throughout the supply chain to ensure the end system is free from safety and security vulnerabilities.

Customization Tools

- Automated application-driven identification and implementation of custom instructions and accelerators is needed in order to customize generic RISC-V design IPs to specific application segments.

Interoperability and Reuse

- There is a need for increased modularity and re-use of subsystems between building blocks. For the design of IP blocks, a gap is that the RISC-V foundation mainly focuses on the ISA interface specification and it has been argued there is relatively little / limited re-use between different components for RISC-V systems.[51]

Hardware/Software Co-certification

- Tool support is needed for co-certification of combined hardware/software systems.

5.9 Roadmap

In this section the important elements of the proposed Roadmap for European Sovereignty in Open Source HW & SW and RISC-V Technologies are given. This is based on the key messages derived from the previous discussions of hardware technology, supporting software and cross cutting requirement needs. These are summarized below:

- Domain-Specific Architectures/Accelerators are one of the major opportunities for advances in computer architecture and have great innovation potential for academia and industry.
- Accelerators need to be integrated with a general-purpose computing environment, in hardware at the chip- and system-levels, as well as into software stacks and programming tools to fully exploit their potential.
- Hardware integration requires significant engineering effort (interface IP blocks, high-performance printed circuit board design) that impedes innovation on accelerators by smaller players, such as the academia and SMEs, who often initiate open source efforts.
- The innovation potential could be unlocked by funding R&D into open standardized interfaces, and their corresponding hardware realizations, as well as scalable & reusable technology templates at both the system-on-chip and computing system-levels. These could then be used by innovators to turn their accelerator architecture ideas into practically usable artefacts (open source releases, products and services offered around the open source releases).
- The challenges of providing software support for innovative accelerators should be addressed. This requires R&D funding for new or improved general purpose accelerator integration software stacks, targeting one or [more] models of computation. Better developer support is also needed for adding new accelerators to established domain specific software frameworks, e.g., machine learning or image processing.
- The existing design sharing and commercialization mechanisms that exist within EUROPRACTICE should be leveraged to promote exchange of developed IP blocks between collaborators and commercialization of IP developed in whole or part within academia. (Note this is already done by the particle physics community to collaborate and contribute designs to CERN).
- Funding should be provided to commission the development/licensing of the required artefacts (e.g., for IP blocks, interfaces and templates) from academia or industry in an open source manner. These should then be made available to users in a "one stop shopping" approach similar to the existing EUROPRACTICE Foundry Access service. Support should also be provided for training & supporting users of the artefacts for a multi-year time period.
- For general-purpose and domain-specific accelerator software stacks a more decentralized approach is needed. Funding should be offered to academia and industry to create/provide the required artefacts in an open source manner, which could then be disseminated using established channels (e.g., github, local repositories). These artefacts may encompass documentation, training materials, and sample back-end implementations for established domain-specific accelerator frameworks. Training and support are also required for a multi-year time period.
- Creating an open source hardware/software ecosystem for innovation in accelerators and domain-specific architectures will require funding schemes which ensure that, after an initial broader experimentation phase, future funds are directed at those technologies that prove most beneficial in practice, e.g., measured in actual user uptake.

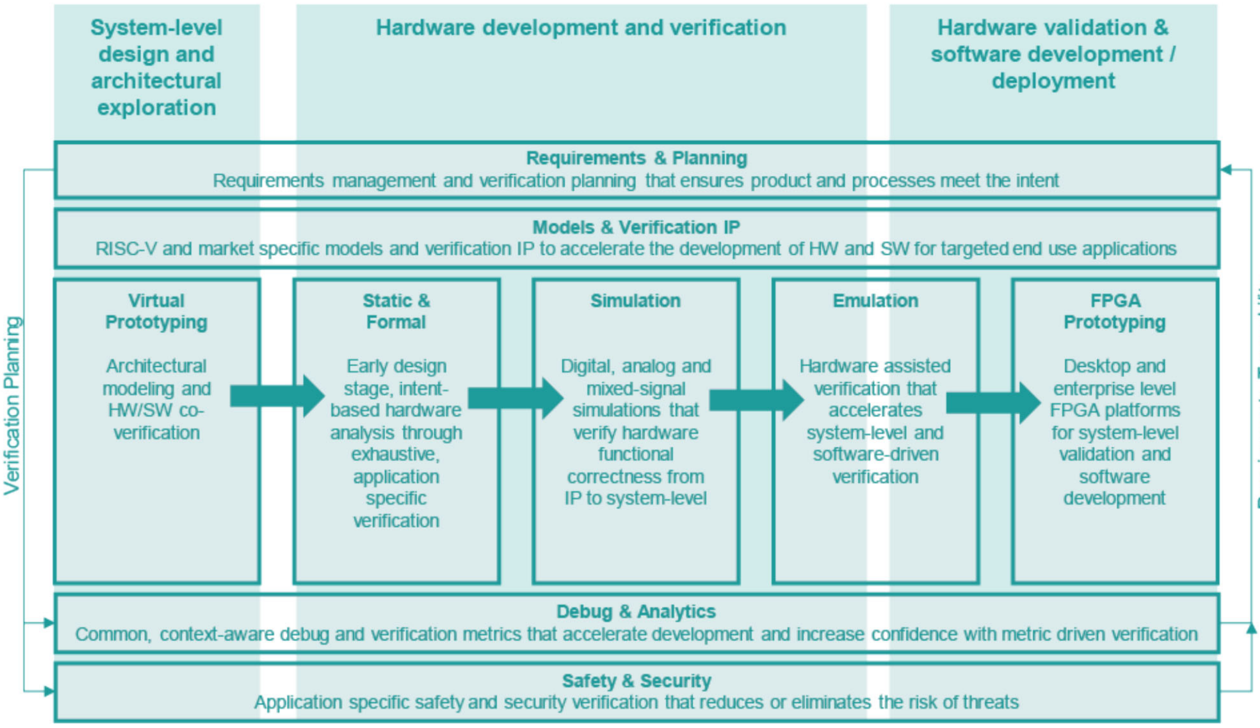


Figure 5.9 - Example Process for RISC-V

In order to meet the needs of the future it is important to sketch out a proposed process that is required to create a working processor. This is done in **Figure 5.9** highlighting the necessary steps and activities in terms of design, fabrication and testing considering the hardware and software aspects. At the application level non-functional properties such as WCET for safety and security need to be addressed. The process shown has been derived for RISC-V, but it could be updated at a later stage for other IP's.

Going beyond this there is a need to develop other IPs. A non-exhaustive list of other IPs identified by the Working Group that will be required in order to give a good coverage of open source cores for making SoCs is given in **Figure 5.9**. These could be used for monolithic SoC integration and for chiplet integration.

Memory controllers addressing Double Data Rate (DDR), Low Power Double Data Rate (LPDDR), High Bandwidth Memory (HBM) and handling the different end use applications and processi
RISC-V based power management controller with placeholder for customer-specific sensors and actuators.
RISC-V based security controller with placeholder for customer-specific secure elements (Physical Unclonable Function (PUF), cryptographic IPs, etc.).
Coherent cache infrastructure for many-cores with controller directory.
Scratchpad memory infrastructure for many-cores with smart DMA.
NoC with on-chip interfaces at router level to connect cores (coherent), memory (cache or not) and IOs (IO coherent or not).
SerDes library in line with PCIe standards.
PCIe controllers and switches.
Chiplet and interposer interfacing units (similar to pads) (only for chiplet + interposer approach).
Trace and debug solutions.
Monitoring IP for safety, security, design integrity.
Guidelines for selection of low/medium speed peripherals (I2C, UART, SPI, etc.).

List of IPs Required

All of these IPs have to be delivered with a verification suite and will need to be maintained constantly to keep up with errata from the field and to incorporate newer requirements.

The availability of standardized metrics is crucial. The application scenario may demand certain boundaries on power, performance, or area of the IP, so searches across different repositories with standardized metrics on these indicators are needed to successfully use these IPs. Building on standardized metrics for all these crucial features, end users will be able to pick the most suited IP for their application and get an idea on needed additional efforts in terms of certifications.

Recommendation - Innovation in open source hardware would profit immensely if lower-level interface blocks could be made available in a low-cost manner, at least for selected chip fabrication processes supported by facilitators such as EUROPRACTICE to allow low-barrier prototyping (e.g., the miniASIC and MPW programs). The availability of automated SoC composition is also desirable to quickly transform innovation into Proof of Concept and to bring productivity gains and shorter time-to-market for industrial projects.

IPs including all artefacts (e.g., source code, documentation, verification suites) should be made available ensuring integrity, authenticity and traceability using certificate-based technologies. Traceability along the supply chain of R&D processes is a foundation for later traceability of supply chains for components in manufacturing and deployment/operation.

Finally, an ecosystem of chiplet + interposers can be enabled via open source development. The key "missing link" here is Die-to-Die (D2D) communication. This is needed to create a chiplet-based design ecosystem. **Development of Open source D2D communication could become one of the most important developments in coming years enabling wide-scale use of chiplets with potential to become a "de-facto" standard.**

5.9 Prioritization of Open Source Ecosystem Needs

In this section the short-, mid- and long-term needs for different items on the roadmap are given. They are based on a detailed analysis of gaps and future needs identified. For some critical aspects, proof-of-concept down to physical implementation or layout is needed. For other aspects, a proof-of-concept at synthesis level (netlist) would be sufficient, e.g., for mapping onto a FPGA.

5.9.1 Repository of RISC-V based Processor Platforms

Short-term needs (2-5 years):

- High-end: Highly customizable Multi-core Out of Order 64-bit open source infrastructure with the associated memory hierarchies (caches&coherency, off-chip) and communication (fast cores to cores, cores to accelerators, cores/accelerators to system). This should be suitable for various instances of processor IP.
- Highly customizable high-end domain-specific cores for high-performance embedded system and/or general-purpose application (link with EuroHPC-call on HPC processors)
- Mid-end: Highly customizable mid-end Open Cores (32-bit) with support for advanced 32-bit ISA and system extensions such as security, scalable vector (Zve), tiny FP (Zfinx) or support for specific arithmetics, DSP, bit-manipulation, scalable and customizable interrupt management such as AIA/ACLINT, e.g., for low latency in critical workloads (TinyML, near-sensor processing, secure IoT stacks).

Mid-term needs (5-10 years):

- High-end: Moving coprocessor functions inside the core by adding new instructions (variable precision, customizable vector, tensor, etc.) with all the necessary software environment support (compiler, libraries, automated HW/SW co-design flows) and validation/verification environment.
- Mid-end: Parametrizable open source soft RISC-V cores, and a range of associated interoperable IPs (e.g., interconnect), in relationship with open source EDA flow (this can lead to a tool generating RTL code according to high-level specifications).

5.9.2 Domain-Specific Processor Features

A verified open source hardware for high-assurance and reliable systems via a whitebox approach.

Short-term needs (2-5 years):

- Provide public artifacts for safety and security by architecture at an initial assurance level.
- Provide public artifacts for safety and security by architecture at a high assurance level, e.g., including formal models.

Mid-term needs (5-10 years):

- Demonstrate how the approach can be carried over to more complex compositional architectures (e.g., multipipeline CPU, a full advanced embedded board or even general-purpose desktop computer mainboard).

5.9.3 Repository of Open Source HW Peripheral Blocks

Create in open source all the generic re-usable elements required to create a complete SoC. All the IPs should be interoperable and composable to build a complete SoC.

Short-term needs (2-5 years):

- High-speed open source memory infrastructure (DDR3+, SDRAM, HBM memories).
- Open source high-performance interfacing: PCIe Gen3+, Ethernet 1+G, and USB 2.0/USB 3++ interfaces and controllers, including process-specific analog PHY (SerDes) components.
- Open source test, trace, debug IP blocks (e.g., JTAG, etc.).
- All the ancillary functions (power management, SPI, I2C, etc.) as open source blocks.
- Support for heterogeneous many-core open source SoCs (e.g., scalable coherent caches, scratchpad memories, w/smart DMA, and a coherency-capable network-on-chip).
- Inter-chip communication links with 2+ Gb/s per pin.
- Template for ASIC SoC and PCB designs, the latter for PCIe-attached accelerators, similar to existing FPGA-based evaluation/prototyping boards, but for plugging in user-provided ASICs.
- Easy and affordable “one stop” access to and support for the IP and technologies developed above.
- Open source repository of peripheral models with different levels of accuracy with their verification artefacts.

Mid-term needs (5-10 years):

- Updates of base building blocks to track emerging standards, possibly also adding support for improvement of 2.5D/3D technologies such as High Bandwidth Memory.
- Support open source infrastructure for easy building of SoC from various sources of IPs, in an as automated way as possible.

Long-term needs (beyond 10 years):

- Updates of base building blocks to track emerging standards and technologies.

5.9.4 Interconnect for Real-Time and Mixed Criticality

Verified open source real-time interconnects between logical units such as cores, memory and other peripherals, regardless of whether they are on the same or different chiplet.

Short-term needs (2-5 years):

- Verified open source real-time interconnect proof-of-concept.
- Verified ready-to-use Design IP blocks for safety-critical interconnects.
- Demonstrate use of interconnects by integrating with existing RISC-V open hardware CPUs, memory controllers, and other relevant resources.
- Verification environment for targeting real-time interconnect, including formal proof of worst-case execution time.

Mid-term needs (5-10 years):

- Verified open source real-time interconnects, including mixed criticality, reliability and security, including industrial-grade verification artefacts.
- Verified ready-to-use Design IP blocks for real-time interconnect for mixed criticality. This interconnect shall give system integrators full control over allocation of available bandwidth to different subjects.
- Demonstrate use of interconnects by integrating with existing RISC-V open hardware CPUs, memory controllers, and other relevant resources.
- Verification environment for target for real-time interconnect, including formal proof of worst-case execution time, including mixed criticality cases.

5.9.5 Interconnect for System Integration

IP can be integrated in a monolithic SoC or can be part of chiplets that are physically interconnected through an interposer, which can also provide services (e.g., routing, power supply). A European ecosystem should emerge from this approach, as it is seen as a major evolution in the chip industry by TSMC for example, and now used by AMD, Intel, Nvidia, etc. Industrials are already working on setting up specifications for “Building an open ecosystem of chiplets for on-package innovations”, like ODSA with its “bunch of wires” and the Universal Chiplet Interconnect Express consortium (The promoter members of UCIe are AMD, Arm, Advanced Semiconductor Engineering, Inc. (ASE), Google Cloud, Intel, Meta, Microsoft, Qualcomm, Samsung, and Taiwan Semiconductor Manufacturing Company)[52].

Short-term needs (2-5 years):

- Verified open source interconnect between chiplets and Interposer.
- Verified open source interconnect for active interposers with parametrization bandwidth, latency, energy and performances.
- All the tools and views to easily use various chiplets and design affordable interposers.

- Tools to help the partitioning of complex designs (Design Space Exploration including chiplet/interposers and 3D technologies).

5.9.6 Domain-Specific Accelerators

Development of flexible hardware/software frameworks for accelerator-intensive System-on-Chips.

Short-term needs (2-5 years):

- Create open source template system-on-chip designs including processors, on-chip/off-chip memory and peripheral interfaces/controllers enabling the easy insertion of domain-specific accelerators.
- Create open source software-stacks (e.g., middleware, drivers, advanced OS integration, programming tools) for interacting with the accelerators from software.
- Create reusable and configurable state-of-the-art open source domain-specific accelerators for common operations, e.g., in the areas of security (crypto), ML/AI, communications, information/signal processing.

Mid-term needs (5-10 years):

- Update SoC templates, software stacks and reusable accelerators to track the state-of-the-art, possibly moving to full support for 2.5D/3D technologies such as multi-process heterogeneous stacks.

Long-term needs (beyond 10 years):

- Update SoC templates, software stacks and reusable accelerators to track the state-of-the-art.
- Interoperability with unconventional acceleration hardware solutions, e.g., neuromorphic computing and quantum computing.

5.9.7 Software

A successful processor requires industry-grade software and tools e.g., compilers/debuggers.

Short-term needs (2-5 years):

- Euro Centralized SW Open Source Consortium.
- Development of and or extension of existing compilers and core libraries with emphasis on:
  - Code density
  - Performance
  - Memory Hierarchy
- Safety-certified open source RTOS and/or hypervisor/separation kernel (e.g., Zephyr, Tock or other hypervisors/separation kernels).
- Open source AI codesign tools, HW/SW optimized, in order to tweak the application on the RISC-V platform.

Mid-term needs (5-10 years):

- Open source ISA Extension design tooling/management (Simulate/Iterate/Create).
- Open source Core Validation Test SW.
- Open source Safety/Security Cert Tooling.
- ISA migration tools.

Long-term needs (beyond 10 years):

- Unconventional acceleration software solutions, e.g., neuromorphic computing.

5.9.8 Methodology and EDA Tools

Development of open source SoC development tools and/or RISC-V specific enhancements of leading-edge closed source tools, including interoperability of open and closed source tools.

Short-term needs (2-5 years):

- Tools for design space and architecture exploration and fast simulation of RISC-V based SoC architectures to optimize processor core configuration, memory hierarchy and HW-SW-partitioning for multiple criteria like performance, latency, power and area (e.g., for integration with model-based systems engineering frameworks), tools for automated system-on-chip composition.
- Tools for automated exploration and implementation of RISC-V application specific customization like custom instructions and tightly coupled accelerators.
- Investment in developing open source interoperability standards to interface between different parts of the ASIC development toolchain (both open and closed), encouraging proprietary vendors to adopt the standard.
- Point investment in open source replacements for specific parts of the flow (synthesis, linting, formatting, simulation, DRC, etc.). Focus on integration and functionality, e.g., in the form of stable and easy-to-use reference flows, but less on state-of-the-art algorithms.
- Open source tools for new, more software-driven verification methodologies, possibly based on next-generation hardware description languages, tested on open source IP.
- Investment into open source Continuous Integration infrastructure for open source IPs/cores, extracting metrics, providing feedback to change requests.
- Open source infrastructure with a well-defined semantic layer enabling to use open source IPs and their associated design flows to train AI/ML models enabling faster, higher quality design and verification.
- Investment studies on using selected open source tools in the flow to realize and evaluate practically relevant sample designs, where innovation can be produced using older, lower-complexity processes.
- Automation support for the composition of complex SoCs (heterogeneous mix of cores and accelerators), including generation of hardware/software interfaces and integration/abstraction with ASIC development toolchain (both open/closed).
- Extending tools and tool flows for full compliance with domain specific standards and regulations like functional safety (e.g., ISO26262, IEC61508, EN50129).

Mid-term needs (5-10 years):

- Enhance open source verification tools towards the state of the art, adding more automation, point tools, interactive debug capabilities, and integration with requirements tracing.
- Extend open source tool flows to also cover non-functional aspects, allowing certifiably secure and/or safe designs.
- Update initial open source tools with more advanced algorithms for better Quality of Result and support for more advanced processes, focusing on those with fabrication capabilities located in Europe.
- Enhancing tools for the earlier phases of the EDA pipeline for high-end complex designs
- Enhancing sections of the complete flow for full automation without human expert intervention in the loop as a "push-button-flow".
- Extend open source tool flow to cover a larger part of the complete EDA pipeline. Focus on achieving a front-to-back open source tool flow capable of handling less complex designs on commodity processes.

Long-term needs (beyond 10 years):

- Complete front-to-back open source EDA tool flow capable of realizing even more complex industrial and academic designs with acceptable results, targeting current fabrication processes.

5.9.9 Domain-Specific Demonstrators

It is of key importance that the RISC-V processor families and related Open Source hardware, software and EDA tools are applied in challenging demonstrators.

Short-term needs (2-5 years):

- Domain specific adaptation of RISC-V based processor solutions for safe, secure and reliable computationally intensive applications, e.g., for automotive, industrial automation, medical applications, etc. Such solutions are expected to address appropriate functional and non-functional, high-performance requirements aiming at realizations on advanced technology, e.g., 16 nm Finfet or below.

Mid-term needs (5-10 years):

- Domain specific adaptation of RISC-V based processor solutions for safe, secure and reliable computationally intensive applications, e.g., for automotive, industrial automation, medical applications, etc. Such solutions are expected to address appropriate functional and non-functional, high-performance requirements aiming at realizations on advanced technology, e.g., 7 nm or below.

Long-term needs (beyond 10 years):

- Domain specific adaptation of RISC-V based processor solutions for safe, secure and reliable computationally intensive applications, e.g., for automotive, industrial automation, medical applications, etc. Such solutions are expected to address appropriate functional and non-functional, high-performance requirements aiming at realizations on advanced technology, e.g., 5 nm or below.

5.10 Recommendations

In the following sections the specific recommendations and needs are broken down.

5.10.1 Specific Recommendations

1. Development of Open Source Hardware
   - Fund the creation of open source or easily accessible low-cost/no-cost, fundamental building blocks (e.g., chip IPs (including processors – RISC V, accelerators, peripherals, data management IPs, debug, etc.), templates for SoCs/chiplets/interposers/PCBs, software frameworks for heterogeneous SoC operation) that reduce the high engineering effort required to practically realize a new hardware design and allow creators from academia and SMEs to focus on their actual innovation.
   - The building blocks should have all the views, software supports (drivers), test and documentation so that they can be easily combined (interoperable) and used (support).
   - Make these building-blocks available for free, or with only minimum financial and administrative overheads, from "one stop shops", e.g., by integrating them into service portfolios of organizations such as OpenHW Group or EUROPRACTICE.
   - Ensure that these building blocks are distributed under an open source license that is adapted to hardware artefacts allowing exploitation by all stakeholders (semiconductor industry, OEMs, SMEs, academy/research, etc.).
2. Community Support
   - Provide a one stop shop model with long-term activities and overall support (e.g., advice for licensing, productization, etc.) for SMEs and start-ups.
   - Encourage the use of standard specifications and standardization efforts when gaps are identified.
3. Development of a Chiplets + Interposer Ecosystem in Europe
   - Encourage an ecosystem of chiplet + interposers via open source development. Die-to-Die (D2D) communication is the "missing link" to leverage the chiplet-based design ecosystem, and its development in open source would enable a wide adoption and could result in a "de-facto" standard.
   - A SoC infrastructure template should be developed that includes communication between IPs, interfaces with memories and the external world, supported by tools that allow the easy integration of new chiplets. A validation suite should be developed allowing the rapid design of a complete SoC from various Open source (or not) IPs, together with a set of "best practices" allowing SMEs, start-ups and industries to "make the step" towards the chiplet+interposer approach.
4. Tools, Validation, Methods and Demonstration
   - The development of industrial strength open source design IPs would benefit greatly from proprietary EDA tools. The licensing of the EDA tools should be made available under appropriate conditions, not at the prohibitive prices paid by commercial end-users, and with permissive licensing of the designed IPs.
   - Funding should be provided to create an open source EDA ecosystem. The effort should be guided by the steps (e.g., logic synthesis, placement, routing) of a typical ASIC EDA flow, and proceed by incrementally replacing one or more closed-source steps with open source tools. To enable the required interoperability between open and closed source tools (similar to the FPGA/ASIC space) open interfaces (e.g., APIs, data formats) between the steps are required to allow open source in and open source out.
   - The development of re-usable verification infrastructure (e.g., IP blocks accompanied by test frameworks) should be supported.

- Industrial demonstrators using Open Source IPs (RISC-V hardware, accelerators and SoC IPs) should be supported to validate the complete chain. If these industrial demonstrators include Printed Circuit Boards, the designs of the PCBs should also be licensed under an appropriate open hardware license. Similarly, software developed for these demonstrators and any documentation should be made available under appropriate software and documentation licenses.

5. Special focus on European Need in Terms of Safety/Security Solutions
    - Research projects creating methods to develop safety- and security-critical open source hardware should be supported. Key aspects are collaboration, documentation, verification and certification in open source communities.
    - A re-usable verification infrastructure (e.g., IP blocks accompanied by test frameworks) for safety/security should be established.
    - Industrial demonstrators using RISC-V hardware, especially in the safety/security area, should be supported to promote hardware/software co-certification for safety and security.

5.10.2 Global Longer-Term Recommendations

In addition to short term goals there are also longer-term goals that need to be supported to create a European critical mass and ecosystem in open source. At the same time there is a need to educate the public and industry on the economic and sovereignty benefits of open source approaches.

1. Non-profit Organisation for Coordinating European Open Source HW IP providers
    - A neutral non-profit organisation could be set up for coordinating European Open Source HW IP providers. The aim of this organisation would be to develop a compliance standard that certifies interoperability and industrial readiness of Open Source HW solutions. The organisation would also orchestrate market specific requirements (e.g., safety and security features for the automotive or industrial automation domains). Although the organisation could be funded by the EU, it may have to be open also for companies outside the EU to avoid the emergence of closed Open Source HW IP clusters around the world.
    - Set-up an IP exchange system between academia and industry (integrated in a one stop shop model). This will encourage new business models for EDA vendors, design & IP houses and IC foundries.
    - Ensure that services that allow the realization of test chips or small production runs will continue to support prototyping and small series manufacturing in Europe. This includes providing all the libraries (PDKs) and support for transforming open sources IPs into silicon chips.
2. Educational measures
    - It is very important to communicate the advantages and reasons for open source hardware to the public and industry in an effective way. This requires development of appropriate pedagogical material and communication campaigns. In particular, it is necessary to explain how open source hardware can have a very positive impact on the economy of the EU, and why it is strategically important in guaranteeing digital sovereignty.
    - To change mind sets the EC should provide incentives to public institutions in the EU, including e.g., universities and research laboratories, so that workers in those institutions contribute more to open hardware developments, with that work appropriately recognized in their career development. Open Hardware should become the default paradigm for all hardware development in publicly financed institutions.
3. Summary Table: Key Topics and Timescales

A roadmap needs concrete actions and timescales in order to become operational. The Working Group has thus developed a detailed list of key topics which need addressing in the short term (2-5 years), medium term (5-10 years) and long term (>10 years). These can be used as input into strategic actions to address the core aspects and needs highlighted in this report. Notably there will also be a need for political or bilateral actions between specific stakeholders to take the roadmap forward.

Topics	Overall Topics	Short Term (2-5 years)	Mid Term (5-10 years)	Long Term (> 10 yea s)	
Open HW Base Building Blocks	Repository for open source HW base building blocks.	Create in open source all the elements required to create a complete SoC: (1) Support the creation of: • High-speed open source memory infrastructure (DDR3+, SDRAM, HBM memories); • Open source High performance interfacing: PCIe Gen3+, Ethernet 1+G, and USB 2.0+ interfaces and controllers, including process-specific analog PHY (SerDes) components; • Open source Test, trace, debug IP blocks (e.g., JTAG, etc.); • All the ancillary functions (power management, SPI, I2C, etc.) as open source blocks. • Remark: All the IPs should be interoperable and composable to build a complete SoC. (2) Create support for heterogeneous many-core open source SoCs (e.g., scalable coherent caches, scratchpad memories, w/ smart DMA, and a coherency-capable network-on-chip). (3) Create open source Inter-chip communication links with 2+ Gb/s per pin. (4) Create open source template ASIC SoC and PCB designs, the latter for PCIe-attached accelerators, similar to existing FPGA-based evaluation/prototyping boards, but for plugging-in user-provided ASICs. (5) Establish easy and affordable “one stop” access to and support for the technologies developed in (1-5).	Support updates of base building blocks to track emerging standards, possibly also adding support for 2.5D/3D technologies such as HBM. Supporting open source infrastructure for easy building of SoC from various sources of IPs, as automated as possible.	Support updates of base building blocks to track emerging standards and technologies.	
Processors	Verified open source hardware for high-assurance systems by whitebox approach.	Provide public artifacts for safety and security by architecture at low assurance level. Provide public artifacts for safety and security by architecture at high assurance level, e.g., including formal models.	Demonstrate how the approach can be carried over to more complex compositional architectures (e.g., multipipeline CPU, a full advanced embedded board or even general purpose desktop computer mainboard).	Based on feedback of previous work.	

Topics	Overall Topics	Short Term (2-5 years)	Mid Term (5-10 years)	Long m (> 10 years)	
	High-end application cores for High-performance embedded system and/or general-purpose application (link with EuroHPC-call on HPC processors).	Develop Multi-core Out of Order 64-bit open source infrastructure with all the near memory communication support (caches) and communication (fast cores to cores and core to accelerators). This should be suitable for various instances of processor IP.	Move coprocessor functions inside the core by adding new instructions (variable precision, vector, tensor, etc.) with full software environment support (compiler, libraries) and validation/verification environment.		
	Open Cores (32-bit) with support for advanced 32-bit ISA and system extensions such as security, vector (Zve), tiny FP (Zfinx), DSP, bit-manipulation, fast interrupts (CLIC) for critical workloads (TinyML, near-sensor processing, secure IoT stacks).		Create parametrizable open source soft RISC-V cores and a range of associated interoperable IPs (e.g. interconnect), supported by open source EDA flow.		
Domain-Specific Accelerators	Flexible Hardware/Software Frameworks for Accelerator-intensive System-on-Chips.	(1) Create open source template system-on-chip designs including processors, on-chip/off-chip memory and peripheral interfaces/controllers enabling the easy insertion of domain-specific accelerators. (2) Create open source software-stacks (e.g., middleware, drivers, advanced OS integration, programming tools) for interacting with the accelerators from software. (3) Create reusable and configurable state-of-the-art open source domain-specific accelerators for common operations, e.g., in the areas of security (crypto), Machine Learning/AI, communications, information processing, etc.	Update SoC templates, software-stacks and reusable accelerators to track the state-of-the-art, possibly moving to full support for 2.5D/3D technologies such as multi-process heterogeneous stacks).	Update SoC templates, software-stacks and reusable accelerators to track the state-of-the-art.	

Topics	Overall Topics	Short Term (2-5 years)	Mid Term (5-10 years)	Long Term (> 10 years)	
Interconnect for al-time and mixed criticality	Verified open source interconnects.	Create a verified open source real-time interconnect PoC Develop verified ready-to-use Design IP blocks for safety-critical interconnects. Demonstrate use of interconnects by integrating with existing RISC-V open hardware CPUs, memory controllers, and other relevant resources. Provide verification environment for real-time interconnect, including formal proof of worst-case execution time.	Create verified open source real-time interconnects, including mixed-criticality and security. Develop verified ready-to-use Design IP blocks for real-time interconnect for mixed criticality. This interconnect shall give system integrators full control over allocation of available bandwidth to different subjects. Demonstrate use of interconnects by integrating with existing RISC-V open hardware CPUs, memory controllers, and other relevant resources. Provide verification environment for target for real-time interconnect, including formal proof of worst-case execution time, including mixed criticality cases.	Based on feedback of previous work.	
Software	Organize: Open Source SW. Develop: Software Tools. Execute: Real Time Operating Systems.	Establish Euro Centralized SW Open Source Consortium. Support compiler projects • Code density • Performance • Memory Hierarchy Create Safety Certified Open Source RTOS (e.g., Zephyr).	Create Open source ISA Extension design tooling/management (Simulate/Iterate/ Create). Create Open source Core validation Test SW. Create Open source Safety/Security Certification Tooling ISA migration tools.	Create Open source AI Codesign tools, HW/SW optimized.	
Methodology & EDA Tools	Open source ASIC development tools, including interoperability of open and closed source tools.	(1) Develop open source interoperability standards to interface between different parts of the ASIC development toolchain (both open and closed), encouraging proprietary vendors to adopt the standard. (2) Identify and target investment in open source replacements for specific parts of the flow (synthesis, linting, formatting, simulation, place and route, Static Timing Analysis, Design Rule Checking etc.). Focus on integration and functionality, e.g., in the form of stable and easy-to-use reference flows, rather than state-of-the-art algorithms.	(1) Enhance open source verification tools towards state of the art, adding more automation, point tools, interactive debug capabilities, and integration with requirements tracing. (2) Extend open source tool flows to also cover non-functional aspects, allowing certifiably secure and/or safe designs. (3) Update initial open source tools with more advanced algorithms for better QoR and support for more advanced processes, focusing on those with fabrication capabilities located in Europe. (4) Extend open source toolflow to cover a larger part of the complete EDA pipeline. Focus on achieving	Complete front-to-back open source EDA tool flow capable of realizing even more complex industrial and academic designs with acceptable results, targeting current fabrication processes.	

Topics	Overall Topics	Short Term (2-5 years)	Mid Term (5-10 years)	Long m (> 10 years)	
		(3) Create Open source tools for new, more software-driven verification methodologies, possibly based on next-generation hardware description languages, tested on open source IPs. (4) Investment in open source CI infrastructure for open source IPs/cores, extracting metrics, providing feedback to change requests. (5) Perform investment studies on using selected open source tools in the flow to realize and evaluate practically relevant sample designs, where innovation can be produced using older, lower-complexity processes.	a front-to-back open source tool flow capable of handling less complex designs on commodity processes.		
	Chiplets & interposers for modular architectures.	Develop an Open source Die-to-Die (D2D) communication interface to leverage the chiplet-based design ecosystem to enable wide usage with target to become a “de-facto” standard.	Develop and promote a SoC infrastructure for “chiplet-based” components, with all the views and tools allowing easy partitioning and use of various chaplets (from various origins). Provide support for active interposers.	Develop support for photonic interposers and evolution of automated toolset environment to increase productivity of developers.	
	Analog IPs.	For SoC, develop analog IPs which are an inherent part of the design solutions (PLL, FLL, DC/DC, ADC/DAC...). This will also help address the resistance of EDA companies to allow open sourcing of Analog IPs.			
	Economics and social aspects of open hardware.	Perform research on economics and social aspects of open hardware.			

5.11 Need for Horizontal And Vertical Activities

It is clear that in order to address the needs of Open Source development within Europe both horizontal and vertical activities are needed. These can be summarized as:

Horizontal Activities

- Fundamental Research
- Capacity Building
- Platform Components

Vertical Activities

- Use Case Scenarios
- Application Platforms
- Platform Usage

Recommendation – Horizontal foundational activities and vertical application driven activities are needed to develop new tools and IP libraries which can be built upon, re-used and exploited in future horizontal and vertical calls, systematically growing a repository of benefit to European stakeholders. In parallel with this there is a need for a strategic "Governance Initiative" to coordinate activities, maintain and promote the repository.

In **Figure 5.11** a proposed vision is given on how these horizontal and vertical activities could be combined to support a European RISC-V and Open Source ecosystem. This requires work on "Foundational Building Blocks" to provide both HW and SW building blocks, along with the development of tools to integrate these building blocks into complex SoC designs. Evolution is needed in both the vertical and horizontal activities. Vertically, the basic foundation infrastructure needs to be expanded with new IP's and different technological dimensions. Horizontally, there is a need for development of three types of processor families, ranging from low- via mid- to high-end generic processors. In addition, there is a need to develop application-specific elements that can be added to the generic processors in order to create dedicated solutions for specific domains, e.g., automotive, industry, telecommunications, health, aerospace, defense, etc. There is also a need for horizontal activities addressing security, verification and certification which can then be exploited in the vertical domains targeting specific implementations.

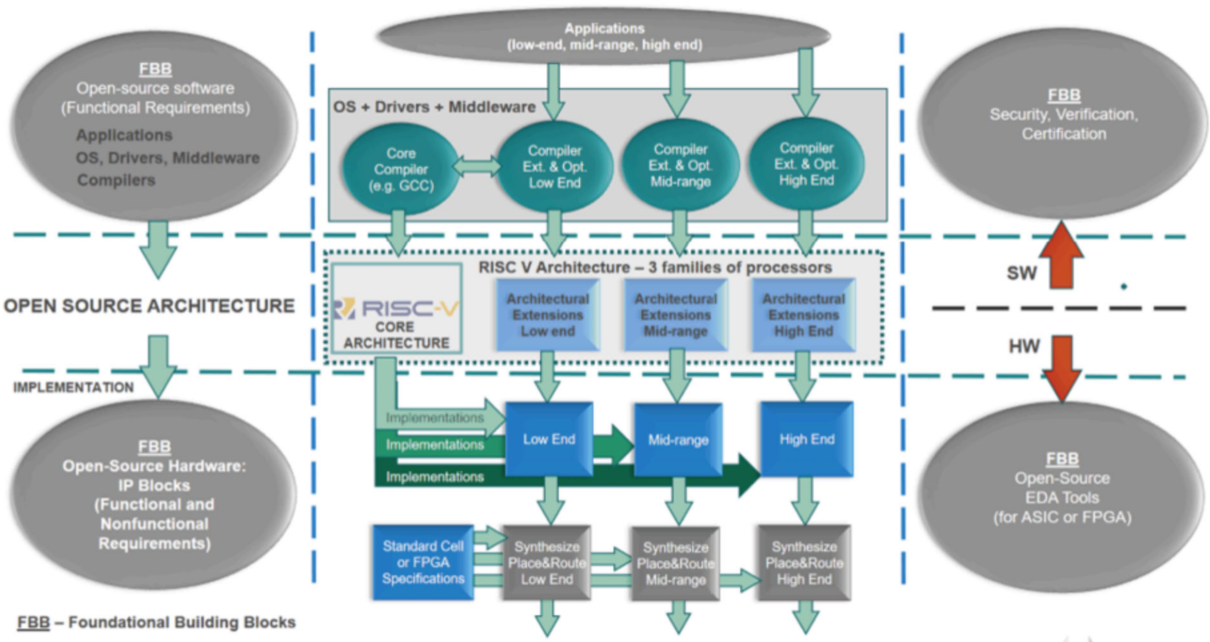
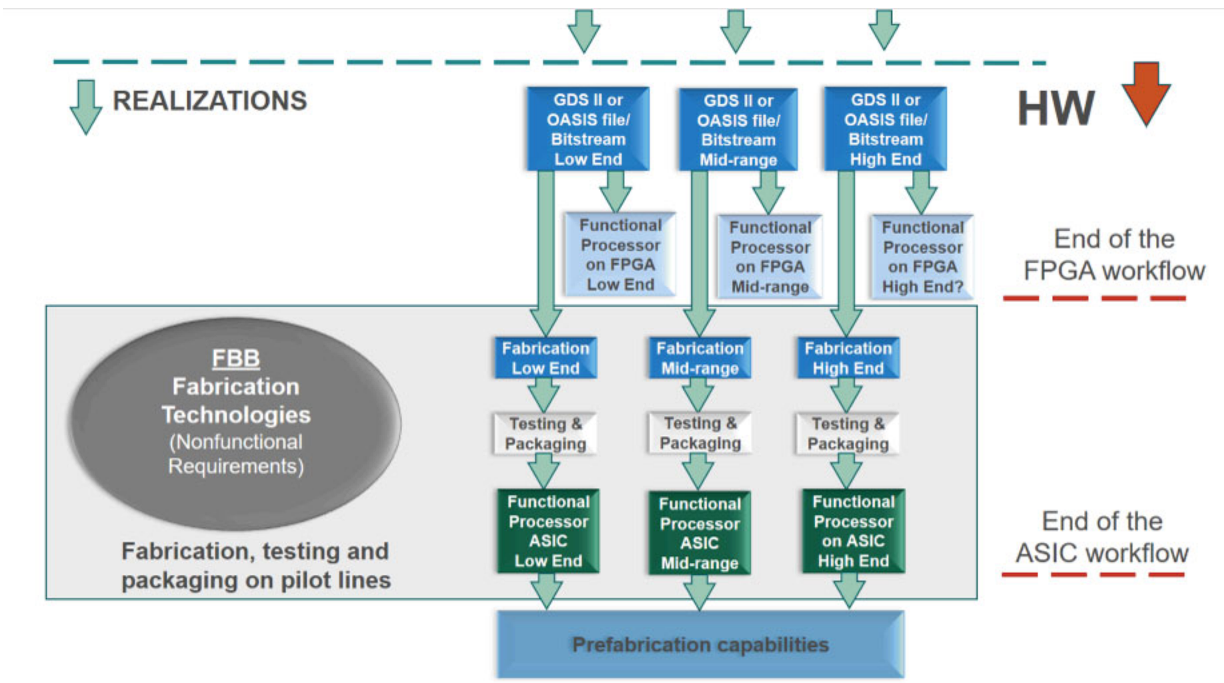


Figure 5.11 - Horizontal Foundational Activities and Vertical Industry Facing Activities

Developing the IP and tools to support the design of new processor architectures will not result in European benefits unless these designs can be physically realized. To support this there is a need for further horizontal and vertical activities as shown on **Figure 5.12** to address the needs of fabrication considering the non-functional properties. Here there is a need for linkage to pilot lines to fabricate, test and package chips, firstly at a prototyping level, and then considering fabrication for specific application sectors.



For the vertical application demonstrators key strategic areas are automotive, industry, high performance computing and communication. A first key vertical application demonstrator could be in the automotive sector in cooperation with OEMs, TIER 1's and TIER 2's. To support development of applications there is a need to intensify interaction with key strategic European stakeholders in application sectors and in other relevant PPPs (e.g., EuroHPC, SNS, etc.) to jointly define specifications, development plans and financing models.

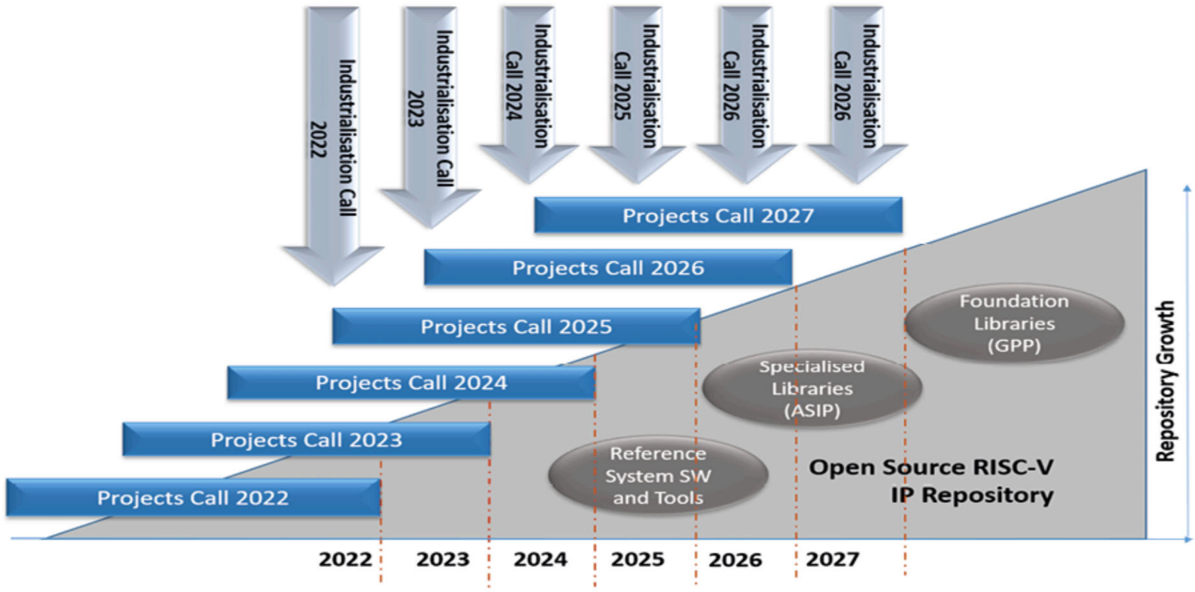


Figure 5.13 - Coordination of Horizonal and Vertical Calls to Create an Open Source RISC-V IP Repository

A possible approach to the coordination of the horizontal and vertical activities is shown in **Figure 5.13** which shows the coordination of project calls addressing both the horizontal aspects as well as application driven vertical calls. In this approach new tools, foundation and specialized IP libraries will be created that are open to the ecosystem. These can be built upon, reused and exploited in future horizontal and vertical calls systematically growing a repository which will be of benefit to European stakeholders.

In parallel with this there is a need for a strategic "**Governance Initiative**" which should be launched as early as possible that orchestrates the EU activities in a top-down manner (to avoid fragmentation), maintains and curates the repository, and promotes the repository to both the open source community and industry. The governance initiative should also define the strategic roadmap of activities, how they interrelate and define expected achievements with milestones. An explicit aim is to ensure multi-year continuity, for instance by reviewing as well as potentially amending submitted research project proposals from a long-time perspective, and by supporting the refinement and definition of future calls.

The governance initiative should be sufficiently supported to be able to technically as well as administratively provide an optimal governance, that takes into account the special needs and characteristics of open source hardware, e.g., with regards to licenses, its economics across different design and production stages, etc. It also can optimize and monitor the submissions, and outcomes of calls, e.g., supervise a study on economic and social aspects of open source hardware innovation. It may additionally have responsibility to maintain the repository and create visibility. The governance initiative could be implemented by an existing organisation, by an agreement between existing organisations, or by a new organisation. The governance initiative should be backed by industry, academia, and users. Members of the working group that wrote this report (i.e., the Open Source Hardware & Software Working Group), are willing to support its initial setup, e.g., by widening the working group and eventually via exchanging with other organisations such as DARPA initiatives on open source hardware.

Already work has begun on developing the Open Source strategy with the launch of the Call 2021-1-IA - Focus Topic 1: Development of open-source RISC-V building blocks with funding of EUR 20 million from Horizon Europe. This addresses mainly horizontal foundational work on RISC-V IP cores, methodologies and tools as well as system software. A further call was launched in 2022 (2022-1-IA- Focus Topic 3) addressing the Design of Customizable and Domain Specific Open-source RISC-V Processors. This call is supported by both national and EUR 20 million of Horizon Europe funding and addresses vertical aspects in terms of specialized processors, e.g., for automotive. The expectation is that the horizontal activities will contribute to the open-source RISC-V IP repository, which is shared and reusable, and the vertical activities will provide domain-specific cores, ready for industrialization which can be either shared or proprietary.

5.12 Concluding Remarks

This report highlights the need for a strong European open source ecosystem to drive competitiveness and enable greater and more agile innovation. There are increasing concerns over security and safety in application markets such as automotive, industrial automation, communications, health and defense where there is a reliance on non-EU developed technologies. Notably open source can be used as a means of retaining sovereignty when the only alternative is to license IPs from non-EU 3_{rd} parties. This is only possible if there is a critical mass of European contributors to open source projects so that a European fork is possible (i.e., create a fully equivalent variant of a given technology) if necessary.

A key message is that there is a need to build or take part in sustainable open source communities such as OpenHW Group, CHIPS Alliance, etc. This is important to get and maintain industrial-grade solutions. Care must be taken to avoid fragmentation by creating too many communities. A challenge is that the current communities are young and deliver limited processor cores and toolchains. There is a need to extend the offer to the community with high-end processors, interconnects, peripherals, accelerators, Operating Systems and SW stacks, Integrated Development Environments (IDE)/Software Development Kit (SDK) with supporting documentation.

A major issue is the availability of Open Source EDA and CAD tools. In addition to open source hardware targeting ASIC implementations there is a need for development tools that do not incur significant licensing costs. Here high-quality open source EDA tools are needed for industrial-grade open source IP cores. A challenge is that Europe has a low footprint in the world of CAD tools, which are critical assets to design and deliver electronics solutions, but there are open source CAD tool initiatives which may bridge this gap.

Considering processors, the involvement in RISC-V International should be strongly encouraged to influence and comply with the "root" specifications. Currently, RISC-V standardizes the instruction set, but additional fields will be needed in the future for more interoperability considering SoC interconnect, chiplet interfaces, turnkey trust and cybersecurity, safety primitives, etc. At a European level it is important to support open source cooperative funded projects to maximize collaborations and create a significant European critical mass to compete with China and the USA. The goal should be to ensure that intellectual property produced is delivered as open source so that European actors can exploit these results.

For open source to be broadly adopted by the European ecosystem, a number of barriers need to be addressed. In particular, there is a perception that building defendable intellectual property is difficult, if not impossible, with open source. This is a key challenge for obtaining financing for open source based start-ups. This requires a redefinition of the business model for many European IP providers, design service providers and public research centers. EU advocacy and financial support is needed for start-ups, SMEs and public research centers to encourage the adoption of open source. At the same time public endorsement of open source and material contribution to open source projects from the leading European semiconductor vendors is critical to provide credibility and momentum to open source at European level. The European design service vendors can also play a key role in supporting the open sourcing of major vendors' IPs, with financial support from the EU.

In order to meet the needs of applications there is a need to address a number of cross cutting issues such as scalability, certification for safety in different application domains, and security. This requires consideration at both the component level and system level.

Key Topics

A number of key topics which need addressing are highlighted in the report. These include "open hardware base building blocks", "verified open source hardware & interconnect" "chiplets and modular interposers", and EDA-tools. More specifically work on open hardware base building blocks should target highlighting benefits of open hardware through proprietary demonstrations (including software, other hardware, mechanical systems) to push the acceptance of open hardware. Considering processors, verified open source hardware should consider leveraging existing open hardware design IP suitable for high-assurance verification and the development of a re-usable verification infrastructure, which includes a complete verification plan considering module level verification, silicon technology specific verification and system integration verification including safety, security and other non-functional aspects (e.g., code quality, clock domain crossing, reset domain crossing, power domain crossing, resource efficiency, power efficiency). This should also consider IP blocks accompanied by verification frameworks for safety/security starting from ESL down to RTL. The aim here would be to release these frameworks and tests under permissive or weakly reciprocal licenses. The use of open source verification technology should be encouraged but leveraging non-open source verification technologies (e.g., simulation tools, formal verification tools, HW emulators) is also possible, as long as the output of the tools can be integrated and checked with open source verification environments. Going a step further to gain acceptance of open source it would be beneficial to demonstrate certification according to relevant safety and security standards with use cases/demonstrations. This should also include verified open source real-time interconnects, considering mixed-criticality and security.

Chiplets and interposers for modular architectures have been highlighted as an opportunity for the semiconductor industry and here there is a need to address the technical challenges such as architecture partitioning, chiplet-to-chiplet interconnect, testability, CAD tools and flows, advanced packaging technologies. Die-to-Die (D2D) communication, in particular, is the "missing link" to leverage the chiplet-based design ecosystem, and its development in open source would enable a wide use of the approach and could become a "de-facto" standard. This needs to target a high-bandwidth, low-latency, low-energy, ultra-short-reach link between dies.

Open source ASIC development tools are a necessary element of the hardware ecosystem, enabling unlimited access to experimentation and lowering barriers to implementing new ideas by all actors, including new players with a more software background. The availability of at least some open source components that could be used for ASIC design will allow for increased automation, new developer-focused workflows and open the door for cloud- and AI-assisted EDA design.

Finally, other important topics that need addressing are the economics and social aspects of open source hardware innovation as the approach is a significant disruptive change from current practice.

To achieve all of these aims a strategy based on addressing both horizontal foundational aspects of open source development and targeted activities addressing vertical sectors is proposed that will contribute to the creation of a European open source repository of IP blocks and tools.

5.13 ANNEX A – Open Source – what, why, how

5.13.1 Benefits of Open Source within the Value Chain

The ecosystem is wide ranging and diverse including the semiconductor industry, verticals and system integrators, SMEs, service providers, CAD tools providers, open source communities, academics and research. It is an area with many opportunities to create innovative start-ups and service offers. The benefits and attraction of adoption of open source depends on the type of actor and their role within the value chain. These can include:

- Creating innovative products with lower costs and access barriers.
- Providing a faster path to innovation and smoother cooperation between actors (academic, research, industry, SME, alliances) as no Non-Disclosure Agreement (NDA) and commercial license need to be negotiated.
- Influencing technical choices and specifications.
- Allowing customization of open source IP to user needs, delivering differentiating products.
- Sharing development costs.
- Reducing risks related to third-party IP (unbalanced commercial relationship, end of maintenance/ discontinued products, export control and trade wars).
- For RISC-V compliant processors, the advantage of the open source SW ecosystem (compilers, debuggers, Operating Systems, frameworks, etc.).
- Building support and design service businesses based on open source IP.
- Conducting research with easier access to digital technologies.
- Using open source material to educate talented students who will fuel future European successes.
- Better auditing of security and safety, ensuring that solutions can be fully audited and checked/verified (e.g., possibility to look for back doors in open source IPs). This is not possible for IPs licensed from 3_{rd} parties.

The Open Source Ecosystem

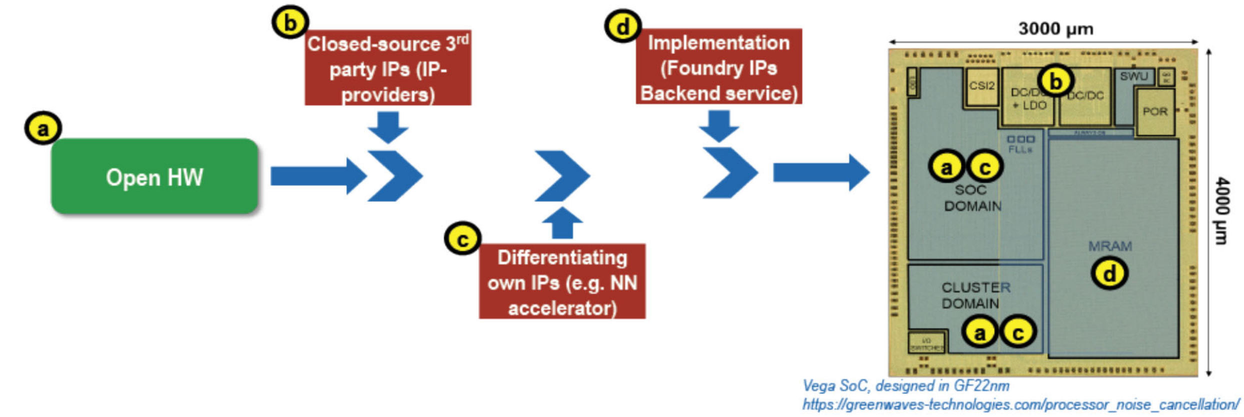


Figure 5.14 - Example Flow of IP Creation of Open Hardware

Open hardware IP is part of an ecosystem composed of other IPs (open source or not), tools to design, simulate, estimate, integrate, synthesize and validate the integration of all IPs (as shown conceptually in Figure 5.14) as well as PDK tools to translate the RTL into a low-level representation specific to the foundry. The foundry utilizes the file generated (generally in GDSII/OASIS format) to produce the masks and the technology process. This is then used to produce a wafer from which the dies are extracted, tested, and packaged, before the final testing and the integration into the final electronic board. At each step significant financial investment is needed.

Although it is generally accepted that high-quality industry ready open source IP that has been through extensive testing, validation and documentation has a lot of value, **it is also believed that there is even greater value from using the RISC-V and Open Source HW/SW IP in the community rather than from the IP itself**. There are different levels of ''Open Source" IP, ranging from standard peripherals to complex application-specific multi-processor-based IP. Business can be made from all types of IP, dependent on the markets addressed, the customers involved and the application needs. Ongoing research, such as by Professor Nagle[53], illustrates the economic benefits of contributing to open source at the state level[54], which would become even more compelling if done at the EU level. The RISC-V ecosystem has the open RISC-V ISA as common starting point for flexibility and interoperability. **The openness of RISC-V also allows inspection which is important for hardware auditability.** Hardware needs to be trusted by consumers and experts alike. This was emphasized by the occurrence of security vulnerabilities such as Meltdown and Spectre which only became visible in 2019 after being deployed in billions of x86 and Arm devices for many years. Open hardware and RISC-V offer a very good chance to improve the situation and the wide acceptance of the RISC-V ISA also allows some degree of interoperability among devices.

5.13.2 Key players in Open Source

The attraction of open source is clear and there is a growing community of developers and contributors to open source repositories both within Europe and at the global scale. A danger is fragmentation of effort and there is a need to strengthen European activities to create a critical mass that can contribute both at the European level, but also on the world stage. This is to ensure that European needs are serviced and that there is more resilience to forks in development that may be deleterious to European interests. Key initiatives identified by the Working Group are described in **Figure 5.14**.

Business can be made on all types of IP, dependent on the markets addressed, the customers involved and the application needs.

RISC-V International is probably one of the most well-known players in open hardware, but it only works on the RISC-V Instruction Set Architecture (ISA) and related specifications. Many entities and individuals have designed RISC-V compatible cores, processors and SoC (https://riscv.org/exchange/cores-socs/). Among them, there are two well-known organizations that host open source co-operations: CHIPS Alliance and OpenHW Group. lowRISC is also notable being mostly recognized for the OpenTitan project.

OpenHW Group is a not-for-profit, global organization driven by its members and individual contributors where hardware and software designers collaborate in the development of open source cores, related IP, tools and software. OpenHW provides an infrastructure for hosting high-quality open source HW developments in line with industry best practices and builds upon the ETH-Zurich/University of Bologna PULP open source project. A good fraction of its 60+ members are from the EU (companies and research centers), some being involved in its governance. The Group hosts the development of several RISC-V cores, CV32E4 embedded family deriving from RI5CY and CVA6 application cores deriving from ARIANE and will extend with TAIGA (from Simon Fraser University) and CV32E20 (compact core deriving from Ibex/Zero-riscy). SoC-level projects are also in the pipeline, as well as several software projects relating to the cores. The list and status of projects can be found at

https://github.com/openhwgroup/core-v-docs/blob/master/program/dashboard/Dashboard.md

CHIPS Alliance is pursuing a vision of creating a completely open hardware ecosystem, with tools, IP, cores, interconnects and more. As a Linux Foundation project it is based in USA but includes parties from all over the world, including EU companies like Antmicro. CHIPS should be viewed as a key strategic partner in pushing the open source technologies to the market. CHIPS Alliance is developing a number of improvements in the open tools workflow, verification, standardization, chiplets and other fields, including:

- SystemVerilog parsers, formatters, linters and other tooling – Verible, Surelog, UHDM, sv-tools (led by Google & Antmicro)
- Open source Universal Verification Methodology (UVM) support in Verilator (led by Antmicro & Western Digital)
- The OpenROAD/OpenLANE and OpenFASoC Application Specific Integrated Circuit (ASIC) design flows (led by Precision Innovation UCSD/UMich, etc.)
- A number of open source Analog tools (with diverse university involvement)
- The RISC-V Domain Validation (DV) verification environment (led by Google)
- The Chisel Hardware Description Language (HDL) and related tools (led by SiFive / Berkeley)
- Advanced Interface Bus (AIB) open source chiplet standard and reference implementation (spearheaded by Intel)
- Field Programmable Gate Array (FPGA) Interchange format, an interoperability standard for open and closed source FPGA tooling which can be used as a reference for similar activities in the ASIC domain (Antmicro / Google)

SkyWater technologies are pursuing a new, open foundry model where their 130nm Process Design Kit (PDK) has been open sourced together with Google, E-Fabless, Antmicro and a number of university partners in a crowdsourcing design platform leveraging also open source IPs and EDA tools. Other PDKs are likely to follow suit. The 3-some partnership offers free access to MPW (Multi Project Wafers) sponsored by Google. SkyWater is considered to be quite US-Centric.

EUROPRACTICE is a highly effective integrated service for Europe, that has been supported by the European Commission for more than 25 years. EUROPRACTICE lowers the entry-barrier for academic institutions and less well-established companies (i.e., start-ups and other SMEs) by offering affordable access to a rich portfolio of industrial-grade design tools and prototyping technologies for customized ASICs, Micro-electromechanical Systems (MEMS) and photonics. The service includes initial advice, training and ongoing support, reduced entry costs and a clear route to chip manufacture and product supply.

FOSSi: The Free and Open Source Silicon (FOSSi) Foundation predates the OpenHW Group and the CHIPS Alliance, with a strong European presence on its Board of Directors. It is completely different in that no big corporate actors are involved in its governance. Its legitimacy in the open source silicon arena is derived from their long-standing commitment to facilitate the sharing of open source digital designs and their related ecosystems. In this regard, they are well known for organizing the ORConf and LatchUp conference series, as well as managing librecores.org, a place on the Web to share HDL code, seen as the evolution of the old opencores.org.

Key Open Source Initiatives

Going forward the Working Group advocates for building upon and consolidation of Open Source Hardware (OSH) communities where there is a significant participation of European actors, e.g., the OpenHW Group, which is a worldwide initiative, and the CHIPS Alliance, which has a strong US footprint. Notably OpenHW is registered in Canada (a more neutral country) and is in the process of creating OpenHW Europe, as a working group hosted by the Eclipse Foundation (registered in Belgium). The OpenHW Group pays strong attention to verification (processes, environment, coverage, etc.) and open source availability of most verification artefacts. There is also a focus on designs written in widespread Hardware Description Languages (HDLs) to ease adoption by many teams and a rule not to include technology subject to export control laws in any jurisdiction. The CHIPS Alliance provides governance, structure and legal assistance including patents and export controls offering the opportunity for collaboration on legal aspects of open hardware.

Key in Europe is EUROPRACTICE and this already offers significant support for open hardware ICs through a design sharing agreement, which is a legal agreement that requires EUROPRACTICE members to share the IP that they have developed with other members. The standard academic license agreements for IC design tools preclude the sharing of IP, so this design sharing agreement is absolutely essential for any serious open source hardware endeavour. **EUROPRACTICE could extend its current offering and organize an IP exchange system, where academic members could submit their IC IP and gain access to all of the member contributed IP by signing a single agreement to join this 'club' rather than concluding a design sharing agreement for each individual piece of IP. Routes to commercialization from this pool would also exist. Furthermore,**

EUROPRACTICE could negotiate with its European foundries (such as GlobalFoundries, STMicroelectronics and X-FAB) to organize low-cost themed design challenges, similar to the Google-Skywater sponsorship. Ideally, this should fuel hardware exchange between innovators (in the form of chiplets), which would require a standard interposer offer on which chiplets can be tested and brought together into a system.

5.13.3 Business Strategies

There are a number of different business models for commercializing Open Source[55]:

- Service and support firms
- Single-vendor open source firms
- Open source distributor firms

These can be driven by different business models or motivations:

- **Aggressive**: customizing Open Source HW to provide differentiation, e.g., with respect to current commercial cores (e.g., Arm) and targeting advanced processing nodes.
- **Cost-sensitive**: using Open Source HW ''as is'' to reduce cost with respect to licensing costs (e.g., from Arm), or with respect to in-house effort, targeting older processing nodes.
- **Efficiency increase**: design priority setting to re-use standard blocks and focus on the application-specific differentiating elements in the SoC-design, in order to get to the market faster.
- **Share development costs:** share costs for pools of open source IP blocks between actors and manage joint projects in communities, such as the OpenHW Group, the Eclipse Foundation, etc.

The creation of value and the path to real chip production differs according to the types of users of the Open Source IP, such as IDMs, IP vendors, fabless companies, foundries, service companies, etc. From the perspective of a private company, one example on how to address Open Source is the following:

- Commit with academic partners to open source some IPs (theirs or their partners') at the latest when they go into production. This provides them with a time advantage where they think it is critical, i.e., for differentiating features, in an industry where competitive advantage that is driven by timing. Those IPs then enrich the Open Source ecosystem.
- Most of the constituents of an SoC are not differentiating. They are basic constituents that can be easily reused from one design to another. Having a rich open source ecosystem allows companies to focus their efforts on the differentiating features.

5.13.3 Licensing models for Open Source

A key aspect of open source business models is the licensing model and Open Hardware projects typically also include extensive documentation and software for simulation, validation and testing.

Permissive licensing is a mechanism whereby the rights holder gives permissions and imposes very few and lightweight obligations in exchange. For example, one obligation might be to explicitly mention the original author in a piece of modified work. The piece of modified work, and any larger work in which it is included, do not need to be released under the same license.

Weakly-reciprocal licensing, whereby a licensee who modifies the work must release the modified version under the same license. This is where reciprocity comes in. The licensee gets a piece of work with a number of permissions, but must share back in exchange, therefore contributing to a virtuous circle of sharing. The "weak" in "weakly reciprocal" refers to the fact that the reciprocity obligations do not extend beyond the initial piece of work that was shared. The original piece of work can be embedded in a larger work and the sources for that larger work do not need to be released under an open source license.

Strongly-reciprocal licensing is a sharing regime whereby the obligations to share back extend to any larger piece of work embedding the originally licensed work.

Documentation Licensing - When considering documentation, the de-facto standard for sharing is the Creative Commons family of licenses. These include CC0 (permissive with no attribution requirement), CC-BY (permissive with an obligation to recognize the original authors) and CC-BY-SA (reciprocal). These three licenses are adequate to share documentation in the context of the open hardware developments.

Key Licensing Approaches

It is important to choose appropriate licenses for each of the released components. Copyright law is "all rights reserved" by default. This means that putting a file somewhere on the web is not enough because the recipients of that file do not have the basic permissions they need to allow copying, modification and publishing of modified versions of that file. Licenses provide the permissions that the rights holder grants to users so they can benefit from and contribute to the open source effort. There are several pros and cons to different licensing approaches. For software, the best choice of license depends on the project. Apache v2 is a modern permissive license which is used in many successful open source software projects. GNU Lesser General Public License (LGPL v3) and Mozilla Public License (MPL2) are also good choices in weakly-reciprocal contexts, and GNU General Public License (GPL v3) is the de-facto standard for strongly reciprocal work. When considering hardware and IC design there are some issues with using open source software licenses. This applies mainly to reciprocal regimes and permissive licensing is a much simpler in practice. As such Apache v2 can be used as a permissive license for hardware projects in the IC design realm with alternatives being Solderpad v2.1, which takes Apache v2 as a basis and modifies some terms to better adapt them to a hardware context. More specifically the CERN Open Hardware License comes in three variants: permissive, weakly-reciprocal and strongly-reciprocal, providing a one stop shop for hardware licensing. The reciprocal variants take into account the specificities of IC design and its legal and commercial environment.

5.13.4 Approaches to licensing in Europe

Past efforts show that, **in order to reap the benefits of open source, it is vitally important to communicate clearly from the beginning the line between open source and proprietary for a given project/initiative (and how they complement in certain cases), as well as to stick to standard commonly agreed practices and definitions within the open source realm**. In practice this is a very challenging endeavor, so it important to agree on some principles. Open source provides a more level playing field for all actors, and this is an integral part of its success. Many people have tried to get one thing without the other and failed as they both go together. There are ways to give industry a competitive edge inside an open source ecosystem, but it should be understood that preferential access to the source for a component automatically makes that component proprietary. Existing licensing terms include:

- **Truly open source** (i.e., a license approved by OSI). This should already ensure it is not discriminatory to the EU or any other state or group of states.
- **Meaningful for Open Hardware**. Some OSI-approved licenses which were drafted for software are difficult to interpret in the hardware case, which can be very confusing and bring legal uncertainty.
- **Permissive or weakly-reciprocal**

Another approach could be to use exclusive licenses for EU projects. However, as the whole world is embracing the use of RISC-V and both open as well as proprietary IP, there is no need for a special or separate license model for EU purposes only. Despite this there is still a need to select the right licenses using European laws. **Licenses need to be carefully selected to avoid potential contamination or being used for restricting usages for European companies**. Export control laws can restrict the use even of Open Source in some cases, for example, in case of not (yet) fully public disclosure. Some countries' export control regulations, such as the United States, may require taking additional steps to ensure that an open source project is satisfying obligations under local regulations[56].

In many cases the value of Open Source IP is questioned by industry for critical designs. **Open source is considered a risk because of lack of available support and because of the fact that it is open towards the whole world, which could make it difficult to differentiate with competitors.** As Europe is lagging behind today in some domains, it can remain a net beneficiary of open source projects for years, provided that it builds the community and the infrastructure capable of leveraging and influencing those projects. EU funding can contribute to this.

There is a fear of giving everything away to competitors. This is based on an incomplete understanding of best practices to combine open source with commercial activity.

The other big family of relevant actors is that of public institutions. In these institutions there is an ongoing debate as to how much of their work should be made open source. The fear that "giving everything away" to competitors will harm the economy is largely based on an incomplete understanding of best practices to combine open source and commercial activity. One reason why groups in public institutions do not share much publicly is the perception that the results of research should benefit exclusively the subset of people who have financed it. For national and international laboratories, one often hears the objection in terms of the problem of offering all our knowledge and technology for free to others in the world. One possible line of reasoning is that this knowledge can then be used by these external actors to harm our economy, by letting 'them' cheaply manufacture what companies in the financing states could have offered for sale, using profit margins to properly pay their employees and reinvest in innovation. This reasoning is incomplete and can lead to decisions which forego important opportunities to benefit our societies. The opening of the RISC-V Instruction Set Architecture (ISA) itself is a good example of a strategic move from a country (the US) in which the open sourcing was crucial in the generation of large benefits for that country. **RISC-V created a new market because it is open and the main beneficiaries of that new market were institutions and companies which were physically and strategically close to the originating university.**

5.14 ANNEX B – Market trends in the chip market

RISC-V and OpenSource HW/SW IP can be applied in strategic SoC designs targeting specific application domains, thanks to an extensible Instruction Set Architecture (ISA). **The value of the global chips market in 2021 was around USD $550 billion.**[57] The bulk of global demand comes today from end-use applications in computing, including PCs and data centre infrastructure (32%), communications, including mobile handsets and network infrastructure (31%), and consumer electronics (12%). **The growth rate is highest, however, in segments previously ruled by analogue and mechanical technology, such as automotive and industrial manufacturing (12% each)**, as shown in **Figure 5.16**[58].

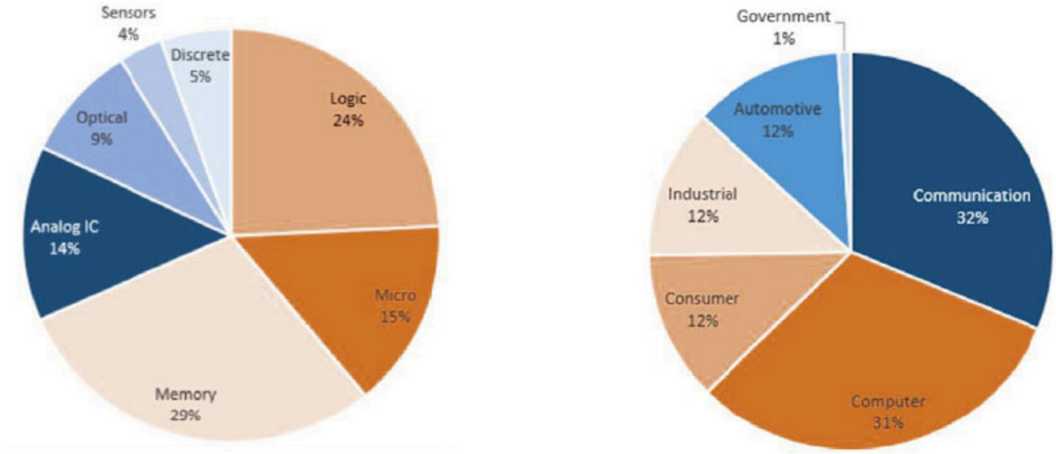


Figure 5.17 - Semiconductor Market Segments, by Device Type and by End-user Sector Demand

Due to strong industry sectors within Europe the split across different domains differs from the worldwide situation with the automotive semiconductors market being 37% of the total European market, the industrial market being 25%, and less of a presence in the computing (15%), communication (15%) and consumer markets (7%)[59]. Overall, the European Semiconductor Industry Association (ESIA) reported that yearly **semiconductor sales in the European market reached USD $ 47.757 Bn in 2021, a 27.3% increase from 2020 and a 27% increase considering the same month in 2020**. At the global scale semiconductor sales in 2021 amounted to USD $ 555.893 Bn, a 26.2% increase from 2020. All figures are based on World Semiconductor Trade Statistics (WSTS) and represent a three-month rolling average unless otherwise indicated. In the table and the graph shown in **Figure 5.18** this growth in European semiconductor sales is clearly shown[60].

Monthly European semiconductor sales

(3-month-average data, except YTD growth which is calculated based on current month data)

Market data for the 3 month moving average ending:								
Region	sales (in billions)		Month on Month growth		Year on Year growth		YTD growth	
	Oct 21	Nov 21	Oct 21	Nov 21	Oct 21	Nov 21	Oct 21	Nov 21
in $:								
Europe	4.140	4.267	2.7%	3.1%	27.8%	26.3%	26.5%	26.4%
Americas	11.027	11.490	2.3%	4.2%	29.7%	28.7%	23.2%	24.4%
Japan	3.890	3.933	1.3%	1.1%	24.7%	19.5%	20.1%	19.3%
Asia Pacific	29.913	29.999	0.3%	0.3%	22.1%	21.7%	25.9%	26.3%
of which China	16.895	16.869	0.2%	-0.2%	22.0%	21.4%	26.1%	26.6%
World	48.971	49.690	1.0%	1.5%	24.4%	23.5%	24.9%	25.3%
In EURO:								
Europe	3.531	3.677	3.3%	4.1%	28.6%	28.4%	19.7%	20.8%
Rate ($/Euro)	1.163	1.142	-1.3%	-3.5%	< Euro against $ versus prev. Year			

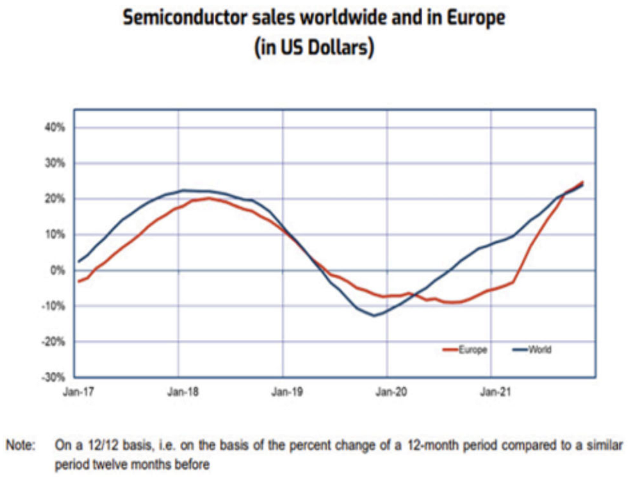


Figure 5.18 - Monthly European Semiconductor Sales and Worldwide Growth in Semiconductor Sales<

The global Microprocessor market was estimated to be almost USD $80Bn in 2020 with the main markets being in computing and communication as shown in **Figure 5.19**. **Europe has a strong position in the embedded market and this rapidly growing market is estimated to account for 21% of sales[61]** .

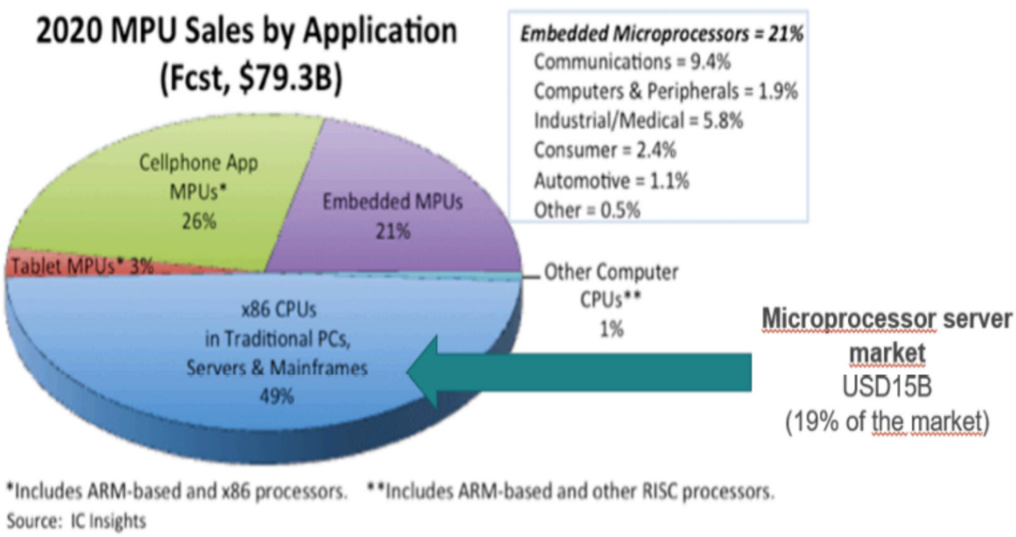


Figure 5.16 - Microprocessor Sales by Application

This is important for Europe as with a global output of more than €1,000 billion, Electronic Systems for embedded and professional applications have overtaken, in production value in 2018, the amount dedicated to the traditional stand-alone electronic goods encompassing smartphones, PCs, TVs, etc. (see **Figure 5.17**[62]).

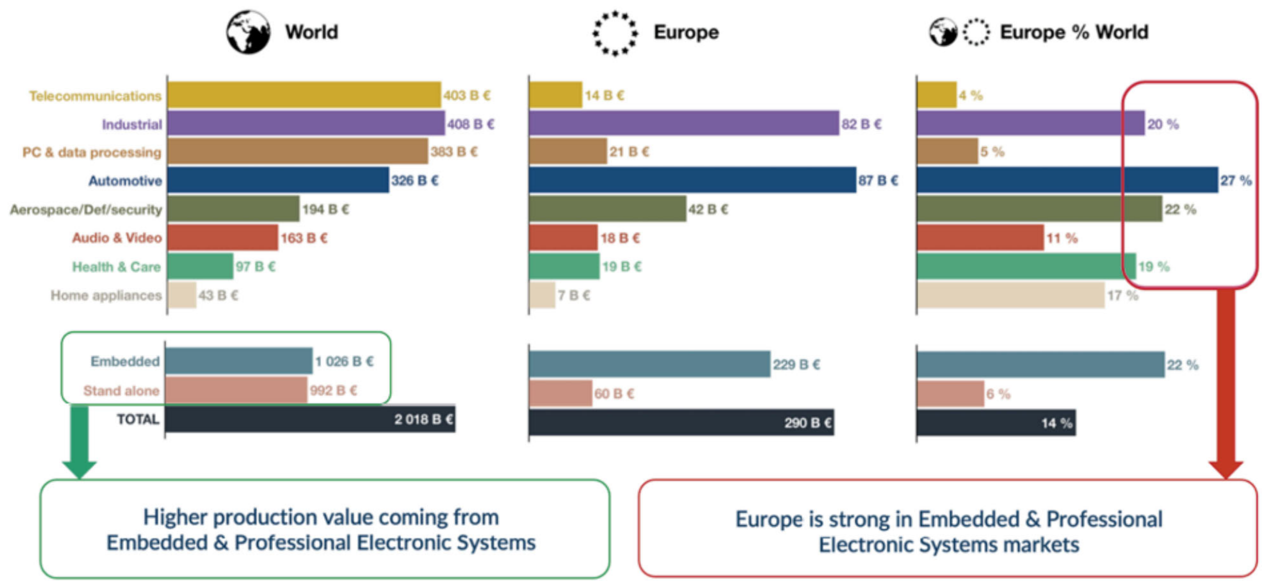


Figure 5.20 - World & Europe Production of Electronics Systems by Application Domain

5.15 Members of the Working Group

Chair:

Patrick Pype NXP Semiconductors

Participants:

Jan Andersson Cobham Gaisler

Luca Benini ETH Zürich / Univ. Bologna

Sven Beyer Siemens

Holger Blasum SYSGO GmbH

Sylvain Depierre Nano-explorer

Marc Duranton CEA

Wolfgang Ecker Infineon

Michael Gielda Antmicro

Edwin Hakkennes Technolution

Andreas Koch Technische Universität Darmstadt

Loic Lietar Green Waves

Andreas Mauderer Bosch

Jan-Hendrik Oetjens Bosch

Jérôme Quévremont Thales

John Round NXP Semiconductors

Javier Serrano CERN

Herbert Taucher Siemens

Jean-Marc Talbot Siemens

EC-Representatives:

Georgi Kuzmanov KDT JU

Panos Tsarchopoulos DG CONNECT

Arian Zwegers DG CONNECT

Editor:

Haydn Thompson THHINK

References

*1 The report was published on **https://digital-strategy.ec.europa.eu/en/library/recommendations-and-roadmap-european-sovereignty-open-source-hardware-software-and-risc-v***

2 In this document CAD tools encompass both classical CAD tools, as well as MCAD, ECAD and system-level tools.

*3 **https://www.eetimes.com/taking-a-risc-expanding-chip-options/#p17***

4 Semico Research Corp., March 2021

5 Semico Research Corp., March 2021

*6 **https://www.counterpointresearch.com/riscv-semiconductor-ip-market-2025/***

*7 **https://www.techdesignforums.com/blog/2020/11/27/risc-v-in-nearly-a-quarter-of-designs-wilson-functional-verification-2020-part-one/?mc_cid=4598aad76e&mc_eid=579aaa7d6f***

*8 **https://blog.westerndigital.com/accelerating-risc-v-ecosystem/***

*9 **https://www.nextplatform.com/2020/08/21/alibaba-on-the-bleeding-edge-of-risc-v-with-xt910/***

*10 **https://www.rolandberger.com/en/Insights/Publications/The-car-will-become-a-computer-on-wheels.html***

*11 Semiconductor Market Size & Share | Industry Growth [2028] (**fortunebusinessinsights.com**)*

*12 MarketsandMarkets. "Avionics Market Size, Growth, Trend and Forecast to 2024." Accessed February 28, 2022. **https://www.marketsandmarkets.com/Market-Reports/commercial-avionic-system-market-138098845.html***

*13 Agirre, Irune, Jaume Abella, Mikel Azkarate-Askasua, and Francisco J. Cazorla. "On the Tailoring of CAST-32A Certification Guidance to Real COTS Multicore Architectures." In 2017 12th IEEE International Symposium on Industrial Embedded Systems (SIES), 1–8. Toulouse: IEEE, 2017. **https://doi.org/10.1109/SIES.2017.7993376**.*

*14 Medical Chip market & Semiconductor Market 2020 report (**electronicsmedia.info**)*

*15 Semiconductor in Healthcare Market Size, Trends, Growth Analysis | 2021 to 2026 (**marketdataforecast.com**)*

*16 VDC Research, Hypervisors, Safe & Secure Operating Systems, 2019. **https://www.vdcresearch.com/Coverage/IoT-Tech/reports/18-Hypervisors-Operating-Systems.html***

17 Knut Blind, Mirko Böhm, Paula Grzegorzewska, Andrew Katz, Sachiko Muto, Sivan Pätsch, Torben Schubert, ''The impact of Open Source Software and Hardware on technological independence, competitiveness and innovation in the EU economy", Final Study Report, Contract Number LC-01346583, SMART 2019/0011.

*18 See **https://www.linuxfoundation.org/resources/publications/understanding-us-export-controls-with-open-source-projects***

*19 **https://cacm.acm.org/magazines/2019/2/234352-a-new-golden-age-for-computer-architecture/fulltext***

*20 **http://nvdla.org/***

*21 **https://github.com/fpgasystems/Vitis_with_100Gbps_TCP-IP***

*22 **https://www.research-collection.ethz.ch/bitstream/handle/20.500.11850/461404/CARRV2020_paper_12_Perotti.pdf?sequence=1&isAllowed=y***

*23 **https://www.athene-center.de/en/research/publications/?tx_vierwdcrisp_publications%5Bpublication%5D=2338&tx_vierwdcrisp_publications%5Baction%5D=show&cHash=aba405ca3a9a90d328778a67ce99e7d2***

*24 **https://riscv.org/wp-content/uploads/2017/12/Tue1212-Customization_of_a_RISC_V_Processor_Zachariasova.pdf***

*25 **https://edge.seas.harvard.edu/files/edge/files/carrv_workshop_submission_2019_camera_ready.pdf***

*26 **https://github.com/esa-tu-darmstadt/tapasco***

*27 **https://tvm.apache.org/***

*28 **https://sampl.cs.washington.edu/tvmconf/slides/2019/E12-Florian-Stock.pdf***

*29 **https://halide-lang.org**/*

*30 **https://arxiv.org/pdf/1810.00952***

*31 **http://groups.csail.mit.edu/commit/papers/2018/frost_workshop.pdf***

*32 **https://mlir.llvm.org/***

*33 **https://sifive.cdn.prismic.io/sifive%2Fcab05224-2df1-4af8-adee-8d9cba3378cd_tilelink-spec-1.8.0.pdf** or **https://bar.eecs.berkeley.edu/projects/tilelink.html***

*34 **http://eembc.org***

35 https://valgrind.org/

36 https://dynamorio.org/

37 Tverdyshev, et al., MILS Architecture, 2013, EURO-MILS, http://dx.doi.org/10.5281/zenodo.45164.

38 https://www.semiconductors.org/strengthening-the-global-semiconductor-supply-chain-in-an-uncertain-era/

39 https://github.com/google/CFU-Playground

40 https://github.com/google/riscv-dv

41 https://github.com/openhwgroup/core-v-verif

42 https://antmicro.com/blog/2021/07/open-source-systemverilog-tools-in-asic-design/#p17_uvm-is-in-the-picture

43 https://docs.opentitan.org/hw/

44 H. Legenvre, P. Kauttu, M. Bos, and R. Khawand, "Is Open Hardware Worthwhile? Learning from Thales' Experience with RISC-V," Research-Technology Management, vol. 63, no. 4, pp. 44–53, Jul. 2020, doi:10.1080/08956308.2020.1762445; Blind, Knut, Mirko Böhm, Paula Grzegorzewska, Andrew Katz, Sachiko Muto, Sivan Pätsch, and Torben Schubert. "The Impact of Open Source Software and Hardware on Technological Independence, Competitiveness and Innovation in the EU Economy," 2021, https://ec.europa.eu/newsroom/dae/redirection/document/79021

45 https://core-v-verif-verification-strategy.readthedocs.io/en/latest/

46 Wilhelm, Reinhard. "Determining Reliable and Precise Execution Time Bounds of Real-Time Software." IT Professional 22, no. 3 (May 1, 2020): 64–69. https://doi.org/10.1109/MITP.2020.2972138.

47 Kästner, Daniel, Markus Pister, Simon Wegener, and Christian Ferdinand. "Obtaining Worst-Case Execution Time Bounds on Modern Microprocessors." Nuremberg, 2018. https://www.absint.com/papers/2018_ew_tw.pdf; Agirre, Irune, Jaume Abella, Mikel Azkarate-Askasua, and Francisco J. Cazorla. "On the Tailoring of CAST-32A Certification Guidance to Real COTS Multicore Architectures." In 2017 12th IEEE International Symposium on Industrial Embedded Systems (SIES), 1–8. Toulouse: IEEE, 2017. https://doi.org/10.1109/SIES.2017.7993376.

48 https://www.sogis.eu/documents/cc/domains/sc/JIL-Composite-product-evaluation-for-Smart-Cards-and-similar-devices-v1.5.1.pdf

49 Ge, Qian, Yuval Yarom, and Gernot Heiser. "No Security Without Time Protection: We Need a New Hardware-Software Contract." Proc. 9th Asia-Pacific Workshop on Systems - APSys, 2018. https://doi.org/10.1145/3265723.3265724.

50 M. Fadiheh, A. Wezel, J. Mueller, J. Bormann, S. Ray, J. Fung, S. Mitra, D. Stoffel, W. Kunz: An Exhaustive Approach to Detecting Transient Execution Side Channels in RTL Designs of Processors, https://arxiv.org/abs/2108.01979

51 Taylor, Michael Bedford. "BaseJump STL: SystemVerilog Needs a Standard Template Library for Hardware Design." In 2018 55th ACM/ESDA/IEEE Design Automation Conference (DAC), 1–6. San Francisco, CA: IEEE, 2018. https://doi.org/10.1109/DAC.2018.8465909.

52 https://www.uciexpress.org/

53 https://www.hbs.edu/faculty/Pages/profile.aspx?facId=566431

54 https://www.hbs.edu/ris/Publication%20Files/19-103_70f212c8-c4fe-4989-ac99-e03cf8bbf02d.pdf

55 Prof. Dr. Dirk Riehle, Univ. Erlangen, Bayave GmbH, „Commercial Open Source", OFA Symposium, 16 Nov. 2021.

56 https://www.linuxfoundation.org/tools/understanding-us-export-controls-with-open-source-projects/

57 Ravi, Sarah. "Global Semiconductor Sales, Units Shipped Reach All-Time Highs in 2021 as Industry Ramps Up Production Amid Shortage." Semiconductor Industry Association, February 14, 2022. https://www.semiconductors.org/global-semiconductor-sales-units-shipped-reach-all-time-highs-in-2021-as-industry-ramps-up-production-amid-shortage/

58 Semiconductor Industry Association Factbook 2021 (figures for 2020)

59 https://www.eusemiconductors.eu/esia

60 ESIA Release Note, 18 February 2022

61 https://epsnews.com/2020/09/14/total-microprocessor-sales-to-edge-slightly-higher-in-2020/

62 "Creating Intelligent Digital Systems Together", A study commissioned by the INSIDE Industry Association (INSIDE-IA), 24 November 2021

6

Strategic Research and Innovation Agenda 2025

APPENDIX B

Acronyms used in the document

2D, 3D	*Two dimension(al), three dimension (al)*
5G	*Fifth-generation communication network*
6G	*Sixth-generation communication network*
A&P	*Assembly and packaging*
AAL	*Ambient Assisted Living*
ACA	*Anisotropic conductive adhesive*
ACES	*Autonomous, connected, electric and shared*
ACK	*Alexa Communication Kit*
ADAS	*Advanced driver-assistance system*
AEB	*Automated Emergency Braking*
AF-EAF	*Air Force Enterprise Architecture Framework*
AFIoT	*Architecture Framework for the Internet of Things*
AFM	*Atomic force microscopy*
AI	*Artificial Intelligence*
AIOTI	*Alliance for the Internet of Things Innovation*
AIoT	*Artificial Intelligence of things*
ALD	*Atomic layer deposition*
AlN	*Aluminum nitride*
ALU	*Arithmetic logic unit*
ANN	*Artificial Neural Network*
AMD	*Age-related macular degeneration*
AMS	*Analogue/mixed signal*
ANC	*Active noise cancellation*

API	*Application programming interface*
AR	*Augmented reality*
AS	*Autonomous system*
ASIC	*Application-specific integrated circuit*
ATE	*Automated test equipment*
AUTOSAR	*AUTomotive Open System Architecture*
B2B	*Business-to-business*
B2C	*Business-to-consumer*
BATX	*Baidu, Alibaba, Tencent and Xiaomi*
BCD	*Bipolar CMOS DMOS*
BCI	*Brain–computer interface*
BDVA	*Big Data Value Association*
BEOL	*Back end of line*
BESS	*Battery Energy Storage System*
BEV	*Battery electric vehicle*
BGA	*Ball grid array*
BiCMOS	*Bipolar CMOS*
BIST	*Built-in self-test*
BOM	*Bill of materials*
BOX	*Buried oxide*
C&K	*Competence and knowledge*
CAD	*Computer-aided design*
CAFCR	*Customer Objectives, Application, Functional, Conceptual and Realisation Model*
CAGR	*Compound annual growth rate*

Cath lab	*Catheterisation laboratory*
CAV	*Connected autonomous vehicle*
CB	*Conductive-bridge*
CBRAM	*Conductive-bridging RAM*
CCAM	*Connected, Cooperative and Automated Mobility*
CDR	*Carbon dioxide removal*
CEAP	*Circular Economy Action Plan*
CFD	*Computational fluid dynamics*
CFET	*Complementary Field Effect Transistor*
CIS	*CMOS Image Sensors*
CMOS	*Complementary metal–oxide–semiconductor*
CMS	*Components, Modules and Systems*
cMUT	*Capacitive micromachined ultrasound transducer*
CNN	*Convolutional neural network*
CNT	*Carbon nanotube*
CPS	*Cyber-physical system*
CPU	*Central processing unit*
CrMMC	*Carbon-reinforced metal matrix composites*
CRM	*Critical Raw Material*
CSA	*Climate Smart Agriculture*
CT	*Computed tomography*
CVD	*Chemical vapour deposition*
D2D	*Device-to-device*
D2W	*Die to wafer*
DCS	*Distributed control systems*
DfA	*Design for assembly*

DfM	*Design for manufacturing*
DfR	*Design for reliability*
DfX	*Design for excellence*
DL	*Deep learning*
DMA	*Direct Memory Access*
DNN	*Deep neural network*
DPP	*Digital Product Password*
DPU	*Data Processing Unit*
DRAM	*Dynamic random access memory*
DSA	*Directed self-assembly*
DSL	*Domain-specific language*
DSS	*Decision-support system*
DT	*Drug-targeted*
DUV	*Deep ultraviolet*
E/E	*Electrical/electronic*
EC-RAM	*Electrochemical RAM*
ECPS	*Embedded and cyber-physical system*
ECS	*Electronic components and systems*
ECSO	*European Cyber Security Organisation*
ECU	*Electronic control unit*
EDA	*Electronic design automation*
EFFRA	*European Factories of the Future Research Association*
eHPC	*Embedded high-performance computing*
EHR	*Electronic health record*
EIP-AGRI	*European Innovation Partnership "Agricultural Productivity and Sustainability"*
EMC	*Electromagnetic compatibility*

EMI	*Electromagnetic interference*
EMR	*Electronic medical record*
EMS	*Energy management systems*
eNVM	*Embedded non-volatile memory*
EOL	*End-of-life*
EP	*Engineering process*
EPD	*Environmental product declaration*
EPI	*European Processor Initiative*
EPLCA	*European Platform on Life Cycle Assessment*
ERP	*Enterprise resource planning*
ERTRAC	*European Road Transport Research Advisory Council*
ESAAF	*European Space Agency Architecture Framework*
ESS	*Electronic smart system*
ETP	*European Technology Platform*
ETP4HPC	*European Technology Platform for High Performance Computing*
EU	*European Union*
EUV	*Extreme ultraviolet*
EV	*Electric vehicle*
FAIR	*Facebook AI Research*
FAIRness	*Findability, accessibility, interoperability and reuse*
FCC	*Federal Communications Commission*
FDSOI	*Fully depleted SOI*
FeFET	*Ferroelectric Field Effect Transistor*
FEM	*Finite element method*
FEOL	*Front end of line*
FET	*(depending on the context) Field Effect Transistor or Future and emerging technologies*

FFT	*Fast Fourier transform*
FinFET	*Fin field-effect transistor*
FLOPS	*(flops or flop/s) Floating point operations per second*
FMEA	*Failure mode and effect analysis*
FMI	*Functional mock-up interface*
FMIS	*Farm management information system*
fMRI	*Functional magnetic resonance imaging*
FMU	*Functional mock-up unit*
FPGA	*Field-programmable gate array*
FTJ	*Ferroelectric tunnel junction*
GAA	*Gate All Around*
GAFAM	*Google, Apple, Facebook, Amazon and Microsoft*
GAMAM	*like GAFAM but with the new name of Facebook (Meta)*
GaAs	*Gallium Arsenide*
GaN	*Gallium Nitride*
GDP	*Gross Domestic Product*
GDPR	*General data protection regulation*
GHG	*Greenhouse gas*
GPS	*Global Positioning System*
GPU	*Graphics processing unit*
GWP	*Global warning potential*
HAD	*Highly automated driving*
HBM	*High bandwidth memory*
HCI	*Human–computer interaction*
HEMT	*High-electron-mobility transistor*
HEV	*Hybrid electric vehicle*

HF	*High-frequency*
HFC	*Hydrofluorocarbons*
HIL	*Hardware-in-the-loop*
HIR	*Heterogeneous Integration Roadmap*
HMI	*Human–machine interface*
HMLV	*High mix low volume*
HPC	*High-performance computing*
HSI	*Human-Systems Integration*
HTA	*Hexagon Tensor Accelerator*
HTF	*Heat transfer fluid*
HVAC	*Heating, ventilation and air conditioning*
HVDC	*High-voltage direct current*
HW	*Hardware*
I/O	*Input/output*
I4.0	*Industry 4.0*

IC	*Integrated circuit*
ICT	*(depending on the context) In-circuit test or Information and communications technology*
IDM	*Integrated device manufacturer*
IEA	*International Energy Agency*
IGBT	*Insulated-gate bipolar transistor*
IIA	*Industrial Internet Architecture*
IIoT	*Industrial IoT*
IIRA	*Industrial Internet Reference Architecture*
IMC	*In Memory Computing*
INCOSE	*International Council on Systems Engineering*

iNEMI	*International Electronics Manufacturing Initiative*
InGaAs	*Indium Gallium Arsenide*
InP	*Indium Phosphide*
IoMT	*Internet of Medical Things*
IoT	*Internet of Things*
IP	*Intellectual property*
IP	*Internet protocol*
IPCEI	*Important project of common European interest*
IPM	*Integrated pest management*
IPSR	*Integrated Photonic Systems Roadmap*
IR	*Infrared*
IRDS	*International Roadmap for Devices and Systems*
ISOC	*Internet Society*
IT	*information technology*
IVD	*in vitro diagnostic*
IXP	*Internet exchange point*
JU	*Joint undertaking*
KDT	*Key Digital Technologies*
KFI	*Key failure indicator*
KPI	*Key performance indicator*
LAE	*Large-area electronics*
LCA	*Lifecycle assessment*
LCP	*Liquid crystal polymers*
LCOE	*Levelised cost of electricity*
LDS	*Laser direct structuring*
LED	*Light Emitting Diode*

LLM	*Large Language Model*
LoC	*Lab-on-a-chip*
LV	*Low voltage*
M2M	*Machine-to-machine*
MaaS	*Manufacturing as a service*
MaaS	*Mobility-as-a-service*
MCM	*Multi-chip module*
MCU	*Microcontroller unit*
MDM	*Multi-dimensional metrology*
MEC	*Multi-access edge computing*
MEC	*Mobile edge computing*
Medtech	*Medical technology*
MEMS	*Micro-electromechanical systems*
MES	*Manufacturing execution system*
MES	*Multi-energy system*
MIL	*Model-in-the-loop*
ML	*Machine learning*
MM-ENS	*Multimodal energy system*
MNBS	*Micro-nano-bio system*
MNS	*Micro-nanosystems*
MODAF	*Ministry of Defence Architecture Framework (UK)*
MOEMS	*Micro-opto-electro-mechanical system*
MOF	*Metal–organic framework*
MOOC	*Massive open online course*
MOSFET	*Metal–oxide–semiconductor field-effect transistor*
MPU	*Microprocessing unit*

MR	*Mixed reality*
MRAM	*Magnetic RAM*
MUT	*Micromachined ultrasonic transducer*
MV	*Medium voltage*
NB	*Narrowband*
NEMS	*Nano-electromechanical systems*
NFV	*Network functions virtualisation*
NFVI	*Network functions virtualisation infrastructure*
NLP	*Natural language processing*
NLU	*Natural language understanding*
NMC	*Near Memory Computing*
NoC	*Network on Chip*
NPU	*Neuromorphic processing unit*
NTN	*Non-terrestrial networks*
NV	*Nitrogen vacancies (in diamond)*
NVM	*Non-volatile memory*
OCT	*Optical coherence tomography*
ODD	*Operational design domain*
OECD	*Organisation for Economic Co-operation and Development*
OEF	*Organisation Environmental Footprint*
OEM	*Original equipment manufacturer*
OLED	*Organic LED*
OOC	*Organ-on-a-chip*
OPV	*Organic Photovoltaics*
OSC	*Oxide semiconductor channel*
OSI	*open systems interconnection*

OSS	*Operations support system*
OT	*Operational technology*
OTA	*Over-the-air*
OxRAM	*Oxide-based RAM*
P2P	*Peer-to-peer*
P4	*Predictive, preventive, personalised, participatory*
PAD	*Productivity-aware design*
PCB	*Printed circuit board*
PCM	*Phase-change memory*
PCRAM	*Phase-change RAM*
PCT	*Product category rules*
PDMS	*Polydimethylsiloxane*
PEALD	*Plasma-enhanced atomic layer deposition*
PEBB	*Power electronics building blocks*
PEF	*Product Environmental Footprint*
PFAS	*Per- and Polyfluorinated Substances*
PFC	*Perfluorocarbons*
PFI	*Physical and functional integration*
PGHD	*Patient-generated health data*
PHM	*Prognostic health management*
PI	*Polyimide*
PIII	*Plasma-immersion ion implantation*
PIC	*Photonic Integrated Circuit*
PLC	*Programmable logic controllers*
PLM	*Product lifestyle management*
PMIC	*Power management integrated circuit*

PMUT	*Piezoelectric micromachined ultrasound transducer*
PNN	*Photonic neural network*
PoC	*Point-of-care*
PoCT	*Point-of-care testing*
PoF	*Physics of failure*
PPAC	*Power, performance, area and cost*
PPE	*Personal protective equipment*
ppm	*Parts per million*
PPP	*Public/private partnership*
PSiP	*Power supply in package*
PTEMM	*Process technologies, equipment, materials and manufacturing*
PV	*Photovoltaics*
PVD	*Physical vapour deposition*
PwrSoC	*Power supply on chip*
PZT	*Lead zirconate titanate*
QIP	*Quantum information processing*
QKD	*Quantum key distribution*
QoS	*Quality of service*
QRSC	*Quality, reliability, safety and cybersecurity*
Qubit	*Quantum bit*
qZSI	*Quasi-impedence source inverter*
R&D	*Research and development*
R&D&I	*Research and development and innovation*
RAG	*Retrieval augmented generation*
RAM	*Random-access memory*
RAMI 4.0	*Reference Architecture Model for Industry 4.0*

ReRAM	*Resistive RAM*
RES	*Renewable energy system*
RF	*Radio frequency*
RFID	*Radio-frequency identification*
RL	*Reinforcement learning*
RNN	*Recursive neural network*
ROHS	*Restriction of Hazardous Substances Directive*
ROI	*Return on investment*
RPA	*Robotic process automation*
RRAM	*Resistive RAM*
RT-PCR	*Real-time reverse transcription polymerase chain reaction*
RTE	*Run-time environment*
RTO	*Research and technology organisation*
RUL	*Remaining useful life*
SaaS	*Software as a service*
SAC	*Tin-silver-copper alloy (SnAgCu)*
SAE	*Society of Automotive Engineers*
SCADA	*Supervisory control and data acquisition*
ScAIN	*Scandium aluminium nitride*
SCM	*Storage class memory*
SCM	*Supply chain management*
SDDS	*Smart drug delivery system*
SDG	*Sustainable Development Goal*
SDK	*Software development kit*
SDN	*Software-defined networking*
SDR	*Software-defined radio*

SEAP	*Strategic Environmental Assessment Plan*
SECAP	*Sustainable Energy and Climate Action Plan*
SEES	*Self-powered electrochemical energy storage system*
SGD	*Speech-generating device*
SiC	*Silicon carbide*
SiGe	*Silicon Germanium (alloy)*
SIL	*Software-in-the-loop*
SiN	*Silicon Nitride*
SiP	*System in a package*
SKC	*Skills, knowledge and competence*
SME	*Small and medium-sized enterprise*
SNN	*Spiking Neural Network*
SoA	*Service-oriented architecture*
SoC	*System on a chip*
SoCPS	*System of cyber-physical systems*
SOI	*Silicon-on-insulator*
SoS	*System of Systems*
SOT	*Spin-orbit torque*
SOTIF	*Safety of Intended Functionality*
SPIRE	*Sustainable Process Industry through Resource and Energy Efficiency*
SQUID	*Superconducting quantum interference device*
SRAM	*Static RAM*
SRGM	*Software reliability growth models*
SRIA	*Strategic Research and Innovation Agenda*
SSI	*Smart systems integration*
STDP	*Spike-timing-dependent plasticity*

STEM	*Science, technology, engineering and mathematics*
STS	*Socio-technical system*
STT	*Spin-transfer torque*
SUMP	*Sustainable Urban Mobility Plan*
SUT	*System-under-test*
SW	*Software*
SWM	*Smart Water Management*
TCM	*Threshold change memory*
TCP	*Transmission control protocol*
TEV	*Through-encapsulant via*
TOPS	*Tera operations per second*
TOU	*Time of use*
TPU	*Tensor processing unit*
TPU	*Thermoplastic Polyurethane*
TRL	*Technology readiness level*
TSMC	*Taiwan Semiconductor Manufacturing Company*
TSN	*Time-sensitive network*
TSO	*Transmission system operator*
TSV	*Through-silicon via*
TV&V	*Testing validation and verification*
UAV	*Unmanned aerial vehicle*
UAV	*Unmanned autonomous vessel*
ULP	*Ultra-low power*
UN	*United Nations*
UPS	*Uninterruptible power supply*
UWB	*Ultra-Wide Band*

UWBG	*Ultra-Wide bandgap*
UXV	*Unmanned vehicle*
V&V	*Verification & validation*
V2G	*Vehicle to grid*
V2X	*Vehicle-to-everything*
VCMA	*Voltage-controlled magnetic anisotropy*
VIL	*Vehicle-in-the-loop*
VIS-NIR	*Visible – Near infrared*
VLSI	*Very large-scale integration*
VOC	*Volatile organic compound*
VR	*Virtual reality*
W2W	*Wafer to wafer*
WBG	*Wide bandgap*
WHO	*World Health Organization*
WLP	*Wafer-level packaging*
WLTP	*Worldwide Harmonised Light Vehicle Test Procedure*
WSN	*Wireless sensors network*
XR	*Extended reality*

Introduction

The scope of the ECS SRIA is very broad and spans many disciplines, each of which has developed a specific understanding of some of the terms used in this report. As a result, the same term can have different meanings for specialists in different ECS domains. This glossary defines some of those terms in an exclusive way to ensure there are no inconsistencies across the various chapters. Although there may be readers that feel uncomfortable with a few of the definitions provided here if they differ from what they commonly mean in their own areas, we feel that developing a common language is important in building a strong and integrated ECS community.

SRIA DEFINITIONS

2D materials: materials that can be obtained in a single or very few monolayers, and whose advantageous properties derived from this fact; as opposed to much thicker 'thin films', and even thicker 'bulk materials'

3D integration: a vertical stack of circuitry or integrated circuits (ICs) for meeting electronic device requirements such as higher performance, increased functionality, lower power consumption, and a smaller footprint. In general, 3D integration is a broad term that includes technologies such as: 3D wafer-level packaging; 2.5D and 3D interposer-based integration; 3D stacked ICs (3D-SICs), monolithic 3D ICs; 3D heterogeneous integration; and 3D systems integration.

3D printing: also known as additive manufacturing, this is the construction of a three-dimensional object from a computer-aided design (CAD) model or digital 3D model. The term “3D printing” can refer to a variety of processes in which materials are deposited, joined or solidified under computer control to create a three-dimensional object, with typically the materials (such as liquid molecules or powder grains being fused together) being added on a layer-by-layer basis.

5G: fifth-generation wireless (5G) is the latest iteration of cellular technology, engineered to greatly increase the speed and responsiveness of wireless networks. With 5G, data transmitted over wireless broadband connections can travel at multi-gigabit speeds, with potential peak speeds as high as 20 gigabits per second (Gbps) by some estimates. These speeds exceed wireline network speeds and offer latency of 1 millisecond (ms) or lower, which is useful for applications that require real-time feedback. 5G will enable a sharp increase in the amount of data transmitted over wireless systems due to more available bandwidth and advanced antenna technology. 5G networks and services will be deployed in stages over the next few years to accommodate the increasing reliance on mobile and internet-enabled devices. Overall, 5G is expected to generate a variety of new applications, uses and business cases as the technology is rolled out.

6G (sixth-generation wireless) is the successor to 5G cellular technology. 6G networks will be able to use higher frequencies than 5G networks and provide substantially higher capacity and much lower latency. Millimeter waves (30 to 300 GHz) and terahertz radiation (300 to 3000 GHz) will most likely be used in 6G. Hundreds of gigabits per second transmission rates will be obtained. One of the important goals of the 6G internet is to support one microsecond latency communications. Artificial Intelligence will be important in designing and optimizing 6G architectures, protocols, and operations.

Advanced Packaging : suite of novel technologies, processes and competences that in a cost-and resource-efficient way allows the physical, electrical and functional integration of any set of technological diverse components required to build an advanced electronic system.

Ambient Assisted Living (AAL): information and communication-based products and services that integrate modern technologies (sensors, microcontrollers, connectivity, secure elements, Artificial Intelligence, etc) into the homes and lives of disabled persons, and vulnerable or older adults. These technologies aim to improve the lives of those facing some of the challenges of ageing, and those who care for older people if they need help. An impact of AAL is also in reducing the costs of health and social care.

Artificial Intelligence (AI): the theory and development of information processing systems able to perform tasks usually requiring human intelligence (such as visual perception, speech recognition, decision-making, and translation between languages) with a certain degree of autonomy. Currently, Large language models, or LLMs have received most attention in AI. They are a type of AI that can mimic human intelligence. They use statistical models to analyze vast amounts of data, learning the patterns and connections between words and phrases.

Augmented reality (AR): an interactive experience of a real-world environment where the objects that reside in the real world are enhanced by computer-generated perceptual information, sometimes across multiple sensory modalities, including visual, auditory, haptic, somatosensory and olfactory.

Autonomous system (AS): performs desired tasks in unstructured environments without continuous human guidance.

Biologic drugs: products that are produced from living organisms or contain components of living organisms. Biologic drugs include a wide variety of products derived from human, animal or microorganisms by using biotechnology. Types of biologic drugs include vaccines, blood, blood components, cells, allergens, genes, tissues and recombinant proteins.

Blockchain: decentralised, chronologically updated database with a consensus mechanism created from a network for the permanent digital securitisation of property rights.

Brain–computer interface (BCI): a direct communication interface between a (biological) brain and a technical (IT- and/or ECS-based) system. A BCI can transfer information in both directions – e.g. enabling the brain to control the technical system or enhancing human perception (such as hearing) with additional information from the technical system (e.g. hearing aid).

Care pathway: the sequence of health and care services a patient receives after entering the care systemduring an episode of care.

Cath lab: examination room in a hospital or clinic with diagnostic imaging equipment used to visualise thearteries of the heart and the chambers of the heart.

Chiplet: A chiplet is a tiny integrated circuit (IC) that contains a well-defined subset of functionality. It is designed to be combined with other chiplets in a "Lego-like" assembly which will have the functionalities of a full-fledged IC. Chiplets are interconnected using silicon, glass, and organic substrates-based interposers, through vias, and micro-bumping to enable high-density, high-bandwidth connections between them. Similar to the System-in-Package (SiP) approach, chiplets allow to build circuit blocks using, for each of them, the most adequate process technology in term of cost, power and performance, while recent progress in interconnect technologies enable to achieve data transfer capabilities between those blocks which are significantly higher than what was previously reachable with SiP technology.

Cloud: the on-demand availability of computer system resources, especially data storage (cloud storage) and computing power, without direct active management by the user. The term is generally used to describe data centres available to many users over the internet (from Wikipedia).

Component: a combination of devices and other elements (such as passives) that fulfil a specific need, such as transduction of a single physical parameter within a well-specified case. A component is not self-contained in all its functions, as it requires the close support of other components for operation (e.g. in data processing, power handling, embedded software).

Computer-aided design (CAD): the use of computers (or workstations) to aid in the creation, modification, analysis or optimisation of a design. CAD software is used to increase the productivity of the designer, improve the quality of design, improve communications through documentation, and to create a database for manufacturing.

Conformal electronics: newly developed microelectronics that overcome the standard planar paradigm and can be prepared on arbitrary curvilinear surface.

Contract-based design: a design methodology where the system, itself as well as its constituents (subsystems, components, modules, etc), are described by contracts that are formalised by specifications of their functional behaviour and properties. This is often given in a “assume-guarantee” format (e.g. for a certain software module a contract could be: “If the other components of the system guarantee the availability of input data at certain, well-defined times and if the hardware platform on which this module is running guarantees the availability of certain processing and memory resources (assumptions), then (guarantee) this module will produce its output within a certain, guaranteed time interval”). In this methodology, a designed system is “correct” if (informally): (i) all assumptions of all constituents are met by guarantees of other constituents; and (ii) the contracts of all constituents together imply the contract of the complete system.

Coopetition: a neologism for the act of cooperating and competing at the same time. Companies that compete in the market with their products might still cooperate on topics that are either pre-competitive or non-product differentiating. Typical examples here are interoperability, standards and development processes.

Cyber-physical system (CPS): an ECS in which a physical artefact is controlled or monitored by algorithms. A CPS is the result of tight intertwined hardware and software components capable of creating a link between the physical world and the digital world, to operate on different spatial and temporal scales, exhibit multiple and distinct behavioural modalities, and interact with each other in ways that depend on the context. Examples of CPS include smart grid, autonomous automobile systems, medical monitoring, industrial control systems, robotics systems and automatic pilot avionics.

Cybersecurity: the protection of information against unauthorised disclosure, transfer, modification or destruction, whether accidental or intentional (IEC 62351-2).

Deep edge: the farthest extreme node where subsystems (sensors, actuators, data loggers) interface with the real world. This node is connected to the cloud, but the connection can be intermittent or absent for long periods of time. The emergence of “tiny machine learning” is based on this premise to enable AI in performance-constrained environments (ultra-low power, limited memory size and calculation power), but always very close to the subsystem.

Deep learning (DL): a special form of machine learning based on artificial neural networks, DL is where the system is able to automatically discover the representations needed for feature detection or classification from raw data. The adjective “deep” in deep learning comes from the use of multiple layers in the network (from Wikipedia).

Deeply embedded software: software that runs on dedicated hardware and not on standard microprocessors. In its simplest form, it is called “firmware”.

Dependability: according to IEC 60050-192:2015, dependability (192-01-22) is the ability of an item to perform as and when required. An item here can be a device, component, module or system. Dependability includes availability (192-01-23), reliability (192-01-24), recoverability (192-01-25), maintainability (192-01-27) and maintenance support performance (192-01-29), and in some cases other characteristics, such as durability (192-01-21), safety and security. A more extensive description of dependability is available from the IEC technical committee on dependability (IEC TC 56).

Development or design tools, development or design frameworks, design flow: design tools are software tools supporting engineers with different tasks during system designs. Ideally, these tools are integrated into frameworks that: (i) provide a uniform user interface to all tools; (ii) “sort” the tools according to the different steps in the design process; and (iii) ensure interoperability between the integrated tools. Regardless of whether the tools used are integrated into a framework or not, the order in which the tools are used is called the “design flow”.

Device: in the context of the SRIA, and if it is not further qualified, a device will designate a “packaged chip”, whether it is a packaged integrated circuit (e.g. system on a chip, memory, processor, microcontroller) or a micro-electromechanical system (MEMS)/micro-opto-electro-mechanical system (MOEMS). A device performs a general electrical, electronic or electrical/electronic-physical transduction role.

Digital infrastructure: foundational services necessary to the IT capabilities of a nation, region, city or organisation.

Digital Product Passport: digital identity of a physical product gathering information about the entire product lifecycle.

Digital twin: a digital replica of a living or non-living physical entity. Digital twin refers to a digital replica of potential and actual physical assets, processes, people, places, systems and devices that can be used for various purposes. The digital representation provides both the elements and the dynamics of how the physical entity operates and "lives" throughout its lifecycle. To be useful in systems engineering, digital twins need to be executable (i.e. engineers must be able to use them in simulations as representatives of the actual physical entity) and/or amendable to formal analysis methods. The more aspects of the physical entity are represented in a digital twin, the more useful it becomes.

Divide and conquer strategy: a strategy in systems engineering where a large problem (i.e. designing and building a complex system or even System of Systems) is iteratively broken down ("divided") into smaller problems (i.e. designing subsystems, modules and components), which are then divided further or solved ("conquered"). The results of each step are then integrated into a solution for the next-level larger problem. Divide and conquer typically leads to hierarchical designs; it is also a strategy well suited for distributed developments within supply chains and platform economies.

Edge computing: a computing paradigm where computation and data storage are close to the location where they are needed, to improve response times, save bandwidth and increase independence. It can also include the gateway between deep edge devices and other edge devices (organised in a federation of devices, see fog computing), or with the cloud (modified from Wikipedia).

Embedded (or edge) high-performance computing: provides supercomputing processing performance in rugged, compact and easily deployable computing architectures optimised to work in harsh environments in the field. Bringing high-performance computing capabilities from data centres to field-deployable applications means reducing space, weight and power absorption, increasing resistance, robustness and reliability while maintaining the same advanced computational performance and energy efficiency. Embedded (or edge) high-performance computing is an enabling technology for many vertical domains, such as autonomous driving, UAV, and security and surveillance systems.

Embedded software: the software that runs on embedded and cyber-physical systems, providing the low-level functionalities required to use the available hardware resources, dedicated operating systems, run-time environments, virtualisation and containerisation platforms, application software, micro-services, etc. Embedded software is specifically conceived to optimally exploit the limited hardware resources of embedded and cyber-physical systems. For deeply embedded software, see the separate definition.

Embedded system: an ECS generated from the combination of a microprocessor(s), GPUs or system on a chip, memory, input/output peripheral devices and embedded software that have a dedicated function within a larger mechanical or electrical system.

Extended reality (XR): refers to all real and virtual combined environments and human–machine interactions generated by computer technology and wearables, where the "X" represents a variable for any current or future spatial computing technologies.

Fog computing: an architecture that uses edge devices to carry out a substantial amount of computation, storage and communication locally, and routed over the internet backbone (from Wikipedia).

Functional safety: the ability of a system or piece of equipment to control recognised hazards to achieve an acceptable level of risk – such as maintaining the required minimum of operation even in case of likely operator errors, hardware failures and environmental changes – to prevent physical injuries or damages to the health of people, either directly or indirectly.

Prosthetics: the branch of medicine or surgery that deals with the production and application of artificial body parts.

Healthcare: the preservation of mental and physical health by preventing or treating illness through services offered by the health profession.

Heterogeneous integration: refers to the integration of separately manufactured components into a higher-level assembly (system in a package) that, in the aggregate, provides enhanced functionality and improved operating characteristics. In this definition, components should be taken to mean any unit, whether individual die, MEMS device, passive component or assembled package or subsystem, that are integrated into a single package. The operating characteristics should also be taken in its broadest meaning to include characteristics such as system-level performance and cost of ownership (from ITRS Assembly & Packaging chapter).

Human Systems Integration (HSI): is a transdisciplinary sociotechnical and management approach of systems engineering (SE) used to ensure that system's technical, organizational, and human elements are appropriately addressed across the whole system lifecycle, service, or enterprise system. HSI considers systems in their operational context together with the necessary interactions between and among their human and technological elements to make them work in harmony and cost effectively, from the early design to disposal.

Industry 4.0: the application of technology to digitally transform how industrial companies operate. These technologies include the industrial Internet of Things (IoT), automation and robotics, simulation, additive manufacturing, and analytics. Industry 4.0 is driven by a need to boost efficiency, become more agile to respond to market unpredictability, improve quality, and to enable new business models.

In silico clinical trials: in silico means performed on a computer or via computer simulation. The term characterises biological experiments carried out entirely on a computer. Although in silico studies represent a relatively new avenue of inquiry, they have begun to be used widely in studies that predict how drugs will interact with the body and with pathogens.

In vitro diagnostics: the technique of performing a given procedure in a controlled environment outside of a living organism. Many experiments in cellular biology are conducted outside of organisms or cells. One of the abiding weaknesses of in vitro experiments is that they fail to replicate the precise cellular conditions of an organism, particularly a microbe.

In vivo clinical trials: experimentation using a whole living organism as opposed to a partial or dead organism. Animal studies and clinical trials are two forms of in vivo research. In vivo testing is often employed over in vitro because it is better suited for observing the overall effects of an experiment on a living subject. Integrated practice unit: Involves a shift from the current siloed organisation by specialty department and discrete service to being organised around the patient's medical condition. Care is delivered by a dedicated multidisciplinary team of clinicians who take responsibility for the full cycle of care for the condition, encompassing outpatient, inpatient, and rehabilitative care, and supporting services (e.g. nutrition, social work, behavioural health). The team measures processes and outcomes as a team not individually, and accepts joint accountability for outcomes and costs.

Integrated circuit: an electronic circuit formed on a small piece of semiconducting material, performing the same function as a larger circuit made from electronic building blocks.

Integrated Photonics: technology that in a microchip embodiment produces systems that detect, generate, transport and process light.

Integration platform: an ECS allowing the integration of different systems, applications and services into a single system. They can be found on all layers of the design hierarchy, ranging from "communication backplanes" in hardware design to "reference architectures" and "middlewares" in system engineering, to distributed service platforms in System of Systems. Integration platforms are an important basis for: (i) standardisation; and (ii) platform-based economies.

Internet of Things (IoT): the set of technologies that bring intelligence to objects, enabling them to communicate with other objects or with other devices. IoT describes the network of physical objects – "things" – that perform functions. For example, with these technologies, billions of sensors embedded in everyday devices can be designed to record, process, store and transfer data, and to interact with other devices or systems that use the network's capabilities.

Interoperability: the capability of computing systems to exchange information that can be understood and used by the receiving system.

Key digital technologies: electronic and photonic components, and the software that defines how they work. These technologies underpin all digital systems, including Artificial Intelligence and the Internet of Things.

Lab-on-a-chip (LOC): a miniaturised device that integrates one or several biological or chemical analysis functions on a single chip (e.g. detecting specific proteins).

Large-area electronics (LAE): electronics fabricated utilising printing and roll-to-roll fabrication methods that, as opposed to integrated circuit technologies, can be used on significantly larger substrates. Inorganic and organic inks and pastes are used for printing conductors and active components such as transistors. Substrates in LAE are typically flexible, such as plastic films or paper, giving rise to the term “flexible electronics”.

Large Language Model (LLM): A large language model is a type of language model notable for its ability to achieve general-purpose language understanding and generation. LLMs acquire these abilities by using massive amounts of data to learn billions of parameters during training. LLMs are artificial neural networks (mainly transformers) and are (pre-)trained using self-supervised learning and semi-supervised learning.

Machine learning: ability for a machine to learn by example without being explicitly programmed to perform the target function. This is one method for implementing Artificial Intelligence.

MEMS, MOEMS, NEMS, MNS, MNBS: micro-electromechanical systems (MEMS) originally referred to miniaturised devices that provided a precise mechanical output (typically a small vertical, horizontal, or rotary displacement) upon an electric excitation (e.g. a microrelay), or vice versa, or an electronic signal from a mechanical excitation (e.g. a microaccelerometer or gyroscope). When the objective of such displacement was to interact with light (e.g. a micromirror), the term “micro-opto-electromechanical systems (MOEMS) was used. Gradually, the transduction domain was extended beyond the mechanical one and chemical and biological mediation were also considered. The overall size of MEMS devices could be in the mm or cm range, the term “micro” referring to the dimension of the device’s internal features to be mastered for the device to be functional. The term “nanoelectromechanical systems” (NEMS) is used when such critical dimension falls back into the nano domain. The terms “microsystem”, “micro-nanosystem” (MNS), or “micro-nano-bio system” (MNBS) were alternatively introduced for those small devices amenable to such generalised transduction principles. This kind of device could be fabricated in principle with different materials, but silicon technologies provided a micromachinable material and a miniaturised technology responsive to the integration of the electronic signal to be conveyed or transduced. MEMS, MOEMS, NEMS, MNS and MNBS are very successful means of interaction between the physical and digital worlds, providing information systems with the means to interact with their environment, sensing it, actuating on it or being powered by it.

Model-based design: where design artefacts (the system, subsystems, component, modules, as well as their connections and the environments in which they will be used), are represented by models that are abstract descriptions of certain aspects of such artefacts (typically, their functional behaviour, timing properties, etc). Ideally, these models are: (i) executable, thus usable in simulation and early verification and validation (V&V); and (ii) detailed enough to be usable in formal analysis and test methods.

Module: ensemble of properly integrated components so that their reunion embodies a definite functionality required for the proper working of a system (e.g. sensing and actuation module, control module, communication module, energy provision module). A module is self-contained in hardware and software, making it interchangeable between systems, and allowing higher abstraction level in systems design.

Molecular biology: study of phenomena in terms of biology molecular (or chemical) interactions. Molecular biology emphasises chemical interactions involved in the replication of DNA, its "transcription": into RNA, and its "translation" into or expression in protein – that is, in the chemical reactions connecting genotype and phenotype.

Open source hardware: the blueprint of hardware artefacts that is (partially) freely available and which anyone can use, modify or enhance (depending on different licences associated with the blueprint).

Open source software: software with source code that is (partially) freely available and which anyone can use, modify or enhance (depending on different open source licensing models existing).

Operational design domain (ODD): comprises the "operating conditions under which a given [...] system or feature thereof is specifically designed to function, including, but not limited to, environmental, geographical, and time-of-day restrictions, and/or the requisite presence or absence of certain [environmental] characteristics" (Surface Vehicle Recommended Practice — Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. SAE: J3016, 2018).

Optical coherence tomography (OCT): a non-invasive imaging test that uses light waves to take cross-section pictures of the retina to help with diagnosis. They also provide treatment guidance for glaucoma and diseases of the retina such as age-related macular degeneration (AMD) and diabetic eye disease.

P4 medicine: a shift in medicine from a reactive to a proactive discipline that is focused on predictive, personalised, preventive and participatory (P4). P4 medicine will be driven by system approaches to disease, emerging technologies and analytical tools.

Patient-generated health data (PGHD): health-related data created, recorded or gathered by or from patients (or family members or other caregivers) to help address a health concern.

Personalised medicine: tailoring of medical treatment for patient cohorts to be treated in a unique manner depending on their health status and previous course of a disease and analysis of personal characteristics.

Photonics: Photonics is the science and technology of light. It encompasses generating, guiding, manipulating, amplifying and detecting light. Photonic components process photons (light) analogous to electronic components processing of electrons. A photonic integrated circuit (PIC) or integrated optical circuit is a microchip containing two or more photonic components which form a functioning circuit.

Plug and play components: component with a specification that facilitates the discovery of a hardware component in a system without the need for physical device configuration or user intervention in resolving resource conflicts.

Point of care: the location at which patient services are delivered (excluding hospital, doctor's office, patient's home).

Point-of-care testing (POCT or bedside testing): performance of clinical laboratory testing at the site of patient care rather than in a laboratory, often by non-laboratorians.

Point of need: new model of having critical data and information when and where it is needed rather than at the point of care. These are diagnostics that can be done anytime, anywhere, for anyone – for instance, as a vital part of managing a chronic disease over time, resulting in improved treatment and patient outcomes.

Power Management Integrated Circuit (PMIC): a number of power management circuits (e.g. conversion, storage, conditioning, communication) integrated onto silicon.

Power Supply in Package (PSiP): a number of integrated power circuits and other related power components (e.g. magnetics, capacitors) enclosed in one single package.

Power Supply on Chip (PwrSoC): an integrated circuit that incorporates multiple power source related building blocks (components/circuits) of an electronic system (conversion, management, storage, conditioning, communication) integrated onto silicon.

Predictive maintenance: techniques designed to help determine the condition of in-service equipment to estimate when maintenance should be performed.

Product lifecycle management (PLM): process of managing the entire lifecycle of a product from inception, through engineering design and manufacture, to service and disposal of manufactured products.

Prognostics (a.k.a. health management): a method that permits the assessment of the reliability of the product (or system) under its application conditions. It predicts the occurrence of an event based on current and future operational and environmental conditions to estimate the time at which a system no longer fulfils its function within desired specifications ("remaining useful life").

Prosthetics: the branch of medicine or surgery that deals with the production and application of artificial body parts.

Quality: in this SRIA, quality is defined as "the degree to which a product meets requirements in specifications that regulate how the product should be designed and manufactured, including environmental stress screening (burn-in) but no other type of testing". In this way, reliability, dependability and cybersecurity, which some readers may have expected to be included under quality, will be treated separately.

Quantum computing: an area of computing focused on developing computer technology based on the principles of quantum theory, which explains the behaviour of energy and material on the atomic and sub-atomic levels. A quantum computer utilises quantum entanglement between qubits to solve a set of computationally complex problems efficiently. The computational power of quantum computers is estimated to grow faster than classical computers in the future.

Quantum sensing: sensor technologies that make use of quantum technology.

Quantum technology: the creation, manipulation and detection of single particle quantum states accurately, enabling the use of quantum superposition and entanglement, where quantum states of several particles cannot be described independently, even when spatially separated. Currently, quantum effects typically require very low temperatures and the use of cryogenic technologies.

Recommender-based (methods and) tools: methods and tools in which the current status of a system under design is analysed and evaluated by design-supporting software, which then gives recommendations to the engineer as to possible further steps and/or options for completing the design, ideally together with an evaluation of the pros and cons for each option.

Reliability: the ability or probability, respectively, of a system or component to function as specified under stated conditions for a specified time (ISO 25010).

Safety (a.k.a. functional safety): freedom from unacceptable risk of physical injury or of damage to the health of people, either directly or indirectly as a result of damage to property or the environment (IEC 61508).

Security of ECS (a.k.a. IT security/cybersecurity): in this SRIA, security of ECS is defined as the prevention of illegal or unwanted penetration, intentional or unintentional interference with the proper and intended operation, or inappropriate access to confidential information. Security is considered to be composed of confidentiality, integrity and availability (ISO 21549-2).

Self-X: in self-X, X stands for adaptation, reconfiguration, etc. Usually in self-reorganising systems the major issue is how to self-reorganise while preserving the key parameters of a system, while being coherent with the initial requirements (e.g. performance, power consumption, real time constraints). Self-adaptation and self-reconfiguration has an enormous potential in many applications.

Smart city: an urban area that uses different types of electronic methods and sensors to collect data. Insights gained from that data are used to manage assets, resources and services efficiently; in return, that data is used improve the operations across the city (from Wikipedia).

Smart drug delivery system (SDDS): an advanced method of drug-targeted (DT) delivery. The smart drug delivered by this system must fulfill the following criteria: (i) increase the doses of delivered drug to the targeted body part of interest (tissue/cells/organs); (ii) not be degraded by any of the body fluids; (iii) diminish side effects by improving the efficacy of drug treatment; (iv) absorption of the delivered drug must cross a biological membrane; and (v) drug is released in appropriate dosages to the body part of interest. SDDS is highly complex and involves an integration of various disciplines, such as biology, chemistry and engineering.

Smart systems integration (SSI): (integrated) smart systems incorporate sensing, actuation and control up to cognitive functions to describe and analyse a situation, and make decisions based on the available data in a predictive or adaptive manner, thereby performing smart actions. The enabling principles of these functions include nanoelectronics, micro-electromechanics, magnetism, photonics, chemistry and radiation. SSI is an assembly of technologies that: build products from components; combine functions in products and systems; connect and network systems to other systems; and, importantly, enable systems to receive and store a “knowledge base” – the software that makes them “smart”.

Stretchable electronics: newly developed microelectronics that overcome the standard planar paradigm and can be prepared on stretchable substrates.

Structural electronics: integrating electronics into the body of structural parts of products of different materials.

System: for the purpose of this SRIA, a system is a set of electronic-based constituents (subsystems, modules and components, realised in hardware, software, or both) that are integrated in a way to together allow the system to perform a desired (set of) function(s).

Note that:

- Due to ECS typically being constructed hierarchically, a (e.g. camera or other sensor) "module" being part of the electronic "system" in an autonomous car might itself be referred to as a "system" when designing it (e.g. while integrating lower-level components to together achieve the "camera function") (see also: system in a package, system on a chip, and others).
- The difference between a "system" (comprising subsystems, modules and components) and a "System of Systems" (also comprising subsystems) is that the constituents of a system are chosen and integrated during design-time (i.e. completely under control of the engineers), while in a System of Systems the constituent (sub)systems are independent and dynamically form (and disband) a System of Systems at run-time.

System in a package (SiP): a number of integrated circuits and other electronics building blocks (e.g. MEMS, antennas) enclosed in one single package.

System on a chip (SoC): an integrated circuit that incorporates multiple building blocks of an electronic system, including processors, memory units, accelerators, and input/output ports, and which covers the complete functionality of an electronic system.

System of Systems (SoS): a collection of independent and distributed embedded and cyber-physical systems dynamically composed to generate a new and more complex system, provided with new functionalities and driven by new goals not present in the constituent embedded and cyber-physical systems individually. An SoS must satisfy five characteristics: operational independence of constituent systems; managerial independence of constituent systems; geographical distribution; emergent behaviour; and evolutionary development processes. A system that does not satisfy these characteristics (specifically the first two) is not considered an SoS.

Teleoperation: teleoperation (or remote operation) indicates operation of a system or machine at a distance. It is similar in meaning to the phrase "remote control" but is usually encountered in research, academia and technical environments. It is most commonly associated with robotics and mobile robots, but can be applied to a whole range of circumstances in which a device or machine is operated by a person from a distance.

Telepresence: the use of virtual reality technology, especially for remote control of machinery or for participation in distant events.

Tracking mode simulation: adapting simulation by respective measurements of the real counterpart.

(Technical) Trustworthiness: having some reasonably well thought-out assurance that the technical realisation of a system is worthy of being trusted to satisfy certain well-specified requirements (e.g. safety, security, reliability, robustness and resilience, ease of use and ease of system administration, and predictable behaviour in the face of adversities, such as high-probability real-time performance).

Value-based healthcare: a healthcare delivery model in which providers, including hospitals and physicians, are paid based on patient health outcomes. Under value-based care agreements, providers are rewarded for helping patients improve their health, reduce the effects and incidence of chronic disease, and live healthier lives in an evidence-based way.

Verification and validation (V&V): independent procedures that are used together for checking that a product, service or system meets requirements and specifications, and that it fulfills its intended purpose. Verification checks whether the development implemented the specified requirements of a product correctly ("are we building the product right"), while validation is a system test checking whether a product can fulfil its intended purpose in a real environment ("are we building the right product?").

Virtual commissioning: the practice of using "virtual" simulation technology to "commission" – design, install or test – control software with a virtual machine model before it is connected to a real system.

Virtual reality (VR): computer technology that makes a person feel like they are somewhere else. It uses software to produce images, sounds and other sensations to create a different place so that the user feels they are really part of this other place. Applications of virtual reality can include entertainment (e.g. video games) and educational purposes (e.g. medical or military training).

Wearables: wearable technology is a category of electronic devices that can be worn as accessories, embedded in clothing, implanted in the user's body, or even tattooed on the skin.

X-in-the-loop: where "X" can be hardware-, software-, models-, systems-, etc. The term is used when testing ECS (or parts of an ECS). The system (e.g. component, module) to be tested is called "system-under-test" (SUT). This SUT is embedded into a testbed (or test environment) that provides the necessary input data (according to a specific test scenario), and which then monitors its outputs, comparing these actual outputs to the expected/specified ones. Within these testbeds, data flow therefore forms a "loop" (from the testbed through the SUT back to the testbed). Depending upon the realisation of the SUT (e.g. as a hardware component/module, software module, simulation model, complete system), different testbeds are needed and the resulting test process is called "hardware-in-the-loop", "software-in-the-loop", etc, or when referred to in a general way "X-in-the-loop".

6.3 Main objectives: An analysis of all major challenges

In this ECS-SRIA the Major Challenges identified by the different chapter teams were analysed and finally merged into Main Common Objectives for the ECS community as shown in the following tables.

In the table, Major Challenges numbering is composed by three digits X.Y.Z, where:

- x is the ECS-SRIA Part (1, 2 or 3);
- y is the ECS-SRIA Chapter (1-4 for Part 1, 1-4 for Part 2 and 1-6 for Part 3);
- z is the Major Challenge number in the specific Chapter.

EC Strat. Targets	Industrial Competitiveness	EU Sovereignty	Sustainability and Green Deal	Digital Age (AI, NGC)
ECS R&I Objectives	**Boost industrial competitiveness through interdisciplinary technology innovations**	**Ensure EU sovereignity through secure, safe and reliable ECS supporting key European application domains**	**Establish and strengthen sustainable and resilient ECS value chains supporting the Green Deal**	**Unleash the full potential of intelligent and autonomous ECS-based systems for the European Digital Era**
	Interdisciplinary technology innovation	**Secure, safe and reliable ECS**	**Sustainable & resilient ECS**	**Intelligent and autonomous ECS**
Major Challenges in ECS Research and Innovation	1.1.1 Advanced computing, in-memory, neuromorphic, photonic and quantum computing concepts	1.1.4 Advanced wafer fab equipment and manufacturing solutions	1.1.6 Sustainable semiconductor manufacturing	1.1.1 Advanced computing, in-memory, neuromorphic, photonic and quantum computing concepts
	1.1.2 Novel sensor, actuation and other devices that enable advanced functionality	1.1.5 Advanced packaging, assembly and test equipment solutions	1.2.4 Sustainability	1.1.2 Novel sensor, actuation and other devices that enable advanced functionality
	1.1.3 Advanced integration solutions	1.2.2 Advanced integration solutions	1.3.5 Support for Sustainability by embedded software	1.1.3 Advanced integration solutions
	1.1.4 Advanced wafer fab equipment and manufacturing solutions	1.2.3 Heterogeneous integration	1.4.2 SoS interoperability	1.2.1 Functionality
	1.1.5 Advanced packaging, assembly and test equipment solutions	1.3.6 Software reliability and trust	1.4.5 SoS monitoring and management	1.3.4 Embedding data analytics and AI
	1.2.1 Functionality	1.4.2 SoS interoperability	2.1.1 Increasing the energy efficiency of computing systems and embedded intelligence	1.4.5 SoS monitoring and management
	1.4.1 Open SoS architecture and infrastructure	1.4.3 Evolvability of SoS composed of embedded and cyber-physical systems	2.1.3 Supporting the Increasing Lifespan of Devices and Systems	2.1.1 Increasing the energy efficiency of computing systems and embedded intelligence
	1.4.2 SoS interoperability	1.4.5 SoS monitoring and management	2.2.1 Strengthening EU connectivity technology portfolio in order to maintain leadership, secure sovereignty and offer an independent supply chain	2.1.4 Ensuring European Sustainability of Embedded Intelligence
	2.1.1 Increasing the energy efficiency of computing systems and embedded intelligence	2.1.4 Ensuring European Sustainability of Embedded Intelligence	2.2.4 Architectures and reference implementations of interoperable, secure, scalable, smart and evolvable IoT and SoS connectivity	2.2.3 Autonomous interoperability translation for communication protocol, data encoding, compression, security and information semantics
	2.2.1 Strengthening EU connectivity technology portfolio in order to maintain leadership, secure sovereignty and offer an independent supply chain	2.2.1 Strengthening EU connectivity technology portfolio in order to maintain leadership, secure sovereignty and offer an independent supply chain	2.3.2 Enabling Sustainable Design for Sustainability	2.2.4 Architectures and reference implementations of interoperable, secure, scalable, smart and evolvable IoT and SoS connectivity
	2.2.2 Investigate innovative connectivity technology (new spectrum or medium) and new approaches to improving existing connectivity technology to maintain the EU's long-term leadership	2.2.4 Architectures and reference implementations of interoperable, secure, scalable, smart and evolvable IoT and SoS connectivity	2.4.4 Ensuring safety and resilience	2.4.5 Human systems integration
	2.2.5 Network virtualisation enabling run-time engineering, deployment and management of edge and cloud network architectures	2.3.1 Enabling cost- and effort-efficient Design and Validation Frameworks for High Quality ECS	3.1.3 Green deal: enable climate and energy optimal mobility	3.1.4 Digitalisation: affordable and safe automated and connected mobility for passengers and freight
	2.3.1 Enabling cost- and effort-efficient Design and Validation Frameworks for High Quality ECS	2.4.1 Ensuring HW quality and reliability	3.2.1 Smart & Efficient - Managing Energy Generation, Conversion, and Storage Systems	3.1.5 Edge2cloud mobility applications: added end-user value in mobility
	2.3.4 Managing diversity	2.4.2 Ensuring dependability in connected software	3.2.3 Future transmission grids	3.1.6 AI enabled engineering tool chain: agile collaborative SDV SW development and SDV as well as ADAS/AD validation
	3.1.1 SDV hardware platforms: modular, scalable, flexible, safe & secure	2.4.3 Ensuring cyber-security and privacy	3.2.4 Achieving Clean, Efficient & Resilient Urban/ Regional Energy Supply	3.2.1 Smart & Efficient - Managing Energy Generation, Conversion, and Storage Systems
	3.1.2 SW platforms for SDV of the future: modular, scalable, re-usable, flexible, safe & secure, supporting edge2cloud applications	2.4.4 Ensuring safety and resilience	3.3.1 Responsive and smart production	3.3.1 Responsive and smart production
	3.2.2 Energy Management from On-Site to Distribution Systems	3.1.1 SDV hardware platforms: modular, scalable, flexible, safe & secure	3.3.2 Sustainable production	3.3.3 Artificial Intelligence in Digital Industry
	3.2.3 Future transmission grids	3.1.2 SW platforms for SDV of the future: modular, scalable, re-usable, flexible, safe & secure, supporting edge2cloud applications	3.4.3 Support the development of home as the central location of the patient, building a more integrated care delivery system	3.3.6 Autonomous systems, robotics
	3.2.5 Cross-Sectional Tasks for Energy System Monitoring & Control	3.1.4 Digitalisation: affordable and safe automated and connected mobility for passengers and freight	3.4.4 Enhance access to personalized and participative treatments for chronic and lifestyle related diseases	3.4.2 Enable the shift to value-based healthcare, enhancing access to 4Ps game changing technologies
	3.4.1 Enable digital health platforms based upon P4 healthcare	3.2.2 Energy Management from On-Site to Distribution Systems	3.4.5 Ensure more healthy life years for an ageing population	3.4.4 Enhance access to personalized and participative treatments for chronic and lifestyle related diseases
	3.4.2 Enable the shift to value-based healthcare, enhancing access to 4Ps game changing technologies	3.2.4 Achieving Clean, Efficient & Resilient Urban/ Regional Energy Supply	3.5.3 Environmental protection and sustainable production	3.4.5 Ensure more healthy life years for an ageing population
	3.4.3 Support the development of home as the central location of the patient, building a more integrated care delivery system	3.4.1 Enable digital health platforms based upon P4 healthcare	3.5.4 Water resource management	3.5.2 Food safety
	3.5.3 Environmental protection and sustainable production	3.4.2 Enable the shift to value-based healthcare, enhancing access to 4Ps game changing technologies	3.5.5 Biodiversity restoration for Ecosystems Resilience, Conservation and Preservation	3.5.3 Environmental protection and sustainable production
	3.5.4 Water resource management	3.4.3 Support the development of home as the central location of the patient, building a more integrated care delivery system	3.6.2 Facilitate empowerment and resilience	3.5.4 Water resource management
	3.5.5 Biodiversity restoration for Ecosystems Resilience, Conservation and Preservation	3.4.5 Ensure more healthy life years for an ageing population	3.6.4 Facilitate supportive infrastructure and a sustainable environments	
	3.6.1 Facilitate individual self-fulfilment	3.5.1 Food Security		
	3.6.2 Facilitate empowerment and resilience	3.5.2 Food Safety		
	3.6.3 Facilitate inclusion and collective safety	3.6.3 Facilitate inclusion and collective safety		
	3.6.4 Facilitate supportive infrastructure and a sustainable environments			

	Ensure engineering support across the entire lifecycle of complex ECS-based systems			
	Lifecycle engineering support			
Major Challenges in ECS Research and Innovation	1.2.1 Functionality	1.3.1 Efficient Engineering of Embedded Software	1.3.2 Continuous integration and deployment	1.3.3 Life-cycle management
	1.3.6 Software reliability and trust	1.3.7 Hardware virtualization for efficient SW engineering	1.4.1 Open SoS architecture and infrastructure	1.4.3 Evolvability of SoS composed of embedded and cyber-physical systems
	1.4.4 SoS integration along the life cycle	1.4.5 SoS monitoring and management	2.1.2 Managing the increasing complexity of systems	2.1.3 Supporting the Increasing Lifespan of Devices and Systems
	2.3.1 Enabling cost- and effort-efficient Design and Validation Frameworks for High Quality ECS	2.3.3 Managing Complexity	2.3.4 Managing diversity	2.4.3 ensuring cyber-security and privacy
	2.4.5 human systems integration	3.1.1 SDV hardware platforms: modular, scalable, flexible, safe & secure	3.1.2 SW platforms for SDV of the future: modular, scalable, re-usable, flexible, safe & secure, supporting edge2cloud applications	3.1.6 AI enabled engineering tool chain: agile collaborative SDV SW development and SDV as well as ADAS/AD validation
	3.2.5 Cross-Sectional Tasks for Energy System Monitoring & Control	3.3.4 Industrial service business, lifecycles, remote operations and teleoperation	3.3.5 Digital twins, mixed or augmented reality, telepresence	3.3.6 Autonomous systems, collaborative robotics