# Strategic Research and Innovation Agenda

NESSI

## Executive Summary

NESSI is the Horizon 2020 European Technology Platform (ETP) dedicated to software, services and data. The domains covered by the ETP act as catalysers for the digital revolution that is deeply transforming Europe, and have huge impacts on the way people work, communicate, organize or entertain themselves. Capitalizing on previous documents produced by NESSI, this Strategic Research and Innovation Agenda sets the scene for the remainder of H2020 and proposes a systematic review of the technological challenges to be addressed to build digital systems that need to be distributed and ubiquitous, trusted and sustainable and exhibit new levels of intelligence. These challenges apply to the infrastructures, the software technologies, and to data and knowledge, and represent new opportunities for innovative services transforming all sectors and introducing new business models.

As software is now the sole key enabling technology transversal to all technical innovations, being embedded in solutions and services or providing smart tools for other domains, it should be promoted by Research and Innovation instruments that recognize its unique nature. Software is not only a ‘game changer’ in transforming and digitizing industry and society, it is also the friendly ‘ghost in the machine that cares’ for human needs. Targeting a true Software Continuum will require dedicated programmes, new cross-cutting initiatives, sectorial innovation and new cooperation with various domains from social sciences to education.

## Contents

# 1. Introduction

NESSI, the Networked European Software and Services Initiative, is the Horizon 2020 European Technology Platform (ETP) dedicated to software, services and data. NESSI provides input to the EU Institutions on research actions and technology matters of particular importance to the software domain itself and important to all software-intensive domains. The overall aim of NESSI is to enable the software and services sector to help vitalize the great potential of the European economy and society. NESSI gathers Partners and Members from all over Europe, from industry, research centres and academia, and engages in close dialogue with the European Commission and other stakeholders on topics of relevance to NESSI, including software technologies, software engineering, security and privacy, cloud computing, data-intensive systems, cyber physical systems, and everything-as-a-service

NESSI takes an active role in addressing future challenges of Europe, by working to ensure that sufficient resources are invested in leading-edge research and innovation in the topics of specific relevance to NESSI. Europe has to turn its outstanding R&I potential, its infrastructure and its technological environment into successful product development and marketable products in order to maintain its competitive edge.

Since the publication of NESSI's last *Strategic and Research Agenda* (SRIA) in 2013, NESSI has been instrumental in shaping the Big Data Value cPPP and setting its strategic agenda. NESSI has published three white papers, on *Software Engineering*, *Cyber Physical Systems* (CPS), and *Security and Privacy*, respectively. All three papers address key emerging opportunities and challenges, from the perspective of software, services and data. NESSI has also provided inputs to DG-Connect on possible orientations on Cloud Computing and Software Engineering, and made specific recommendations for the ICT Work Programme 2018 – 2020, focussing on the key role that generic software engineering principles, techniques and tools will play. NESSI argues for key software engineering challenges to be addressed by research and innovation actions towards realizing a Software Continuum. To avoid replicating investments and research effort, NESSI advocates setting up centralized activities on generic software engineering principles, techniques and tools, which may feed complementary domain-specific activities.

As the Digital Single Market strategy unfolds, and as Horizon 2020 continues to launch calls for collaborative projects involving software, services and data, this new SRIA aims at giving a global view of the challenges identified by NESSI, and putting them in a consistent perspective. In anticipation of the preparation of the next Framework Programme, it also challenges the current domain structure and advocates a renewal of Europe's ambition on software and data, commensurate with the expected importance of digitalization of Europe in the next decade.

As a key part of its Strategic Research and Innovation Agenda, NESSI advocates a strong value chain spanning novel research through innovation, adaptation, and entrepreneurship. Whether carried out by academia or industry, large companies or SMEs, research within the areas of relevance to NESSI is a catalyst for innovation related to solving grand challenges in general, and for expanding and supporting software innovation in particular.

## 2. Major trends for the digital economy

As processes, products, services and things continue to be digitized and interconnected on a growing scale, the major transformation towards the Digital Information Society and Economy is accelerating. This transformation spans most industry sectors and its technical trends are rapidly generalized, leading to less time between the emergence of new concepts, solutions, and their adoption. Examples include:

- the commoditization of cloud and big data infrastructures that makes them available to a larger number of enterprises, either on public platforms or as internal solutions;
- the renaissance of artificial intelligence, machine learning and cognitive technologies that allow global players to develop algorithms to gain deeper knowledge and provide innovative services at large-scale;
- the introduction of new architectures and patterns to manage complexity and scale, natively supporting distribution, scalability and resilience (cloud management, software defined databases, micro-services and APIs, …);
- the continuing consolidation of digital hardware providers (phones, network equipment, computers, appliances), along with a shift to software-defined architectures and products, creating new opportunities for synergies and economy of scale;
- the availability of miniaturized hardware for computing and connectivity, along with new standards, increasingly transforming the interaction between the digital world and the physical world through the Internet of Things, integrating sensors and actuators, machine-to-machine communication, fog computing, and ubiquitous personal devices; and
- the evolution of software and product engineering that benefit from automation, the availability of new open-source components and frameworks, and changes in the development lifecycles that introduce more agility while reducing risks and increasing quality.

These positive trends in transformation and digitalization at a large scale also introduce some effects that may require higher level of adaptation or mitigation, such as:

- increased security and privacy risk assessment, as the commoditisation of basic components, the availability of information and knowledge, and the growth of online communities facilitate cyber-crime;
- higher impact of digital failures in densely-interconnected ecosystems of services and infrastructures that form the basis of critical social and economic processes;
- environmental gains resulting from digitization and optimization may be negated by the explosion of digital devices and the complexity and size of algorithms and data; and
- social and economic disruptions as ICT radically transforms industry and services leading to major changes in business models, obsolescence of some skills, automation of some jobs, and changes in behaviours.

Together, these trends lead to a continuous flow of opportunity, based on software and service innovation, to build a better Digital Information Society and Economy, as introduced in the previous NESSI SRIA.

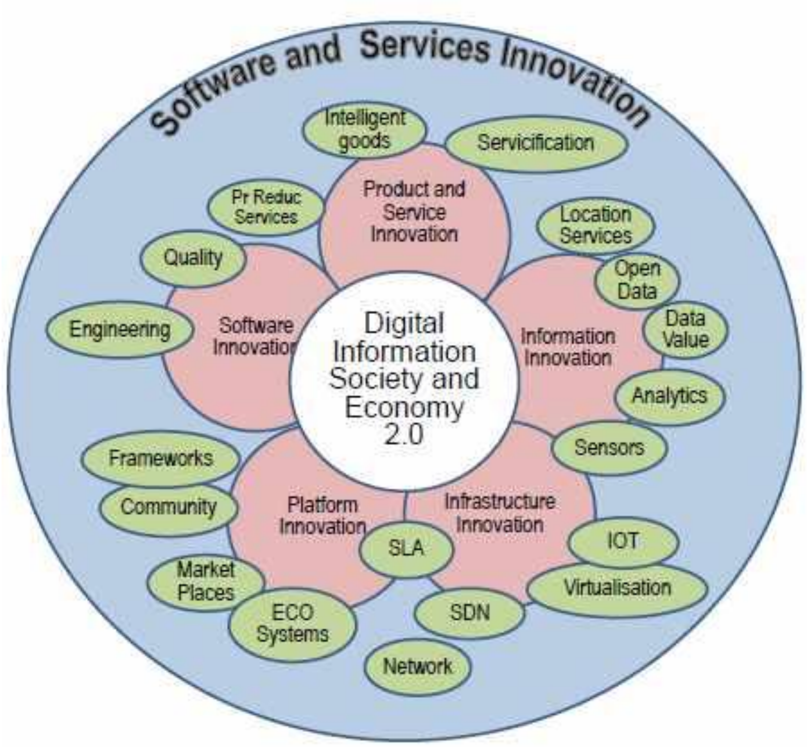



Figure 2.1: Innovations for the Digital Information Society and Economy

On one hand, software innovation is needed in its own right in the software science domain (left top bubble of Figure 2.1) in order to ensure that general software engineering methods and technologies keep up with the ever increasing complexities and challenges that need to be handled by software. On the other hand, the innovations in the other areas (from product and service innovation to platform innovations) will continue be software-driven to a very large extend. As technology and applications continue to conquer new sectors, their synergies create new innovations and domains, or revive older ones - for example, artificial intelligence, cognitive computing, smart agents and robots.

Software will be a key asset to realizing this vision of a digital society and economy, connecting everything together and managing everything. Tremendous challenges remain to be solved in software science to efficiently solve increasingly vast complexities, for instance:

- huge heterogeneity in infrastructure, communication, policies, SLAs;
- security, privacy and trust;
- non-functional aspects such as robustness, availability, recoverability, performance;
- distributed processing, and the trade-off between autonomy and global control;
- dynamic evolution (AI, machine learning, self-adaptation);
- end-user interaction and end-user programming; and
- cyber physical systems eventually ‘merging’ humans and machines.

We see a mix of emerging challenges that are due to the differentiating characteristics of new technology, together with pressing issues from unsolved ‘old’ challenges, which - given the scale and the nature of the digital society - implies that new and better solutions are urgently needed. The future digital society, with its open and dynamic ecosystems, will require the same level of security, safety and trust that the current state of the art in software engineering able us to ensure in dedicated and closed safety-critical systems. Totally new innovations in software methods and techniques are needed to reach the vision of the digital society and to be able to trust its realisation.

# 3. Research and Innovation Challenges

Software, services and data now support the unprecedented process of digitalizing all aspects of modern societies. The digital transformation will spawn the fourth industrial revolution, mixing algorithms, data, physical objects, and biological systems. Software is the key enabler for such mixing, and the glue for interoperation.

In a digital world, software is everywhere: in devices and smart objects, in networks, in servers. Software is not only the programs that describe the behaviours of applications or components but it encompasses any formal description that is understood and used by a computer. A data description is software, a business rule is software, a network configuration file is software, a security assertion is software, text orders given to a chat bot in natural language is software, an SLA with KPIs is software; software is anything that can be analysed, transformed or executed automatically.

## Dimensions of Digital Systems

To better classify the research and innovation challenges in the scope of NESSI, we consider the four dimensions that characterize and will continue to characterize digital systems of all kinds**:** Digital Infrastructures, Software Technologies, Data and Services.

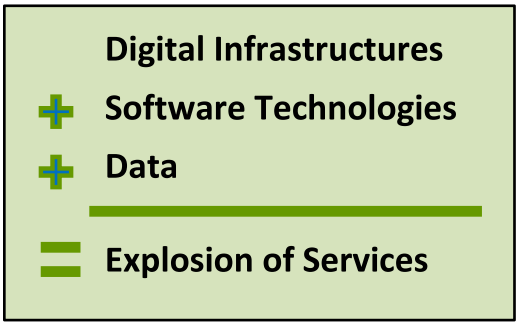



Figure 3.1: Equation of Digital Innovation

### Digital Infrastructures

All digital systems run on some kind of infrastructure, from very small processors embedded in IoT objects to very large data centres, including all personal devices used by most European citizens and the equipment of the communication networks or the utilities' networks. In this SRIA, we consider that infrastructures include the hardware resources of digital systems, the basic layers such as operating systems or virtualisation kernels, and the generic middleware that support the software layers implementing the applications and services. The Internet, the access networks, and cloud computing solutions are examples of such digital infrastructures that now support a major part of all modern societies.

Infrastructures are under a continuous transformation toward convergence and commoditization. The challenge is to deliver proven, sustainable and efficient architectures and solutions for all sectors and services. We shift from vertically separated worlds (networks, CPS, information systems,…) to the emergence of a global consistent 'platform' supporting an ever-expanding variety of use cases, whose characteristics and behaviours are determined by software and data, providing services to all parties (humans, organisations, artefacts).

### Software Technologies

All digital systems rely on software, which provides an unprecedented opportunity for convergence and capacity to evolve for all kind of products, services and processes. Software is everywhere, and software artefacts can take myriad forms. Software technologies are now fully integrated in most sectors as a key enabling technology. Hence, software engineering must cover the full life cycle of numerous products and services, from ideas to maintenance operations, and is a core part of system engineering and process engineering. This involves an ever growing number of actors whose skills have now to include some kind of formalization, modelling or programming (possibly using domain-specific languages).

### Data

Data is the essence of digitalization as it supports the exchange and persistence of state, information, knowledge, rules, etc. Once limited to the so-called data centres, data is now pervasive in all sectors and throughout business, science and society, from the pockets of users to the core of objects, from sensors spread in nature to petabytes in bunkers.

Managing data at all scales, tracing its evolution over time, linking it to the physical world and human beings, making sense of it for beneficial societal impact or monetisation (or both) relies on sound technologies and proven/secure solutions to prevent a divergence between data and reality, misuses of information, or false insights for decisions.

### Services

Services are the way digital solutions fulfil the expectation of actors, either individuals or organisations. Services are dynamically delivered in real-time, delegated to autonomous processes running in the background of society, or embedded in objects to define their behaviours. Some services are explicitly used by humans to accomplish specific tasks, others are invisibly woven in all aspects of digitally empowered sectors; some are consumed with a clear business model, others are provided as part of large ecosystems.

Services are the touch-points where digital systems meet the perception and cognition of human beings and the richness and complexity of human organisation and process. As digital systems becomes smarter, reaching levels of cognition and self-* management, the design and utilisation of services has to take into account multi-disciplinary approaches, including sociology, psychology and economy.

## Characteristics of Computing Systems

Using the above four Dimensions (Digital Infrastructures, Software Technologies, Data and Services), we will address the research and innovation challenges in the scope of NESSI according to three major characteristics of modern computing systems: Physical/Embedded/ Distributed Computing, Secure/Trustworthy/Sustainable Computing, and Smart/Cognitive Computing.

### Physical/Embedded/Distributed Computing

Computing power associated with the relevant data is now available in an ever growing number of objects and systems. It materializes in the miniaturization of processors, memories and sensors, in the digital signal processing associated with modern communication technologies, in the association of processing, engineering and design to create smart objects or smart robots, and the fact that processors are now natively able to communicate to each other. All these factors contribute to building a digital world where computing is performed in a very large variety of contexts and where programs synchronize at local and global scale.

Cyber physical systems and distributed systems have long provided difficult challenges for the computing research and engineering community. Variety in delays and performances, mixes of synchronous and asynchronous behaviours, and strong constraints and requirements are characteristics that make it hard to implement good consistency/optimisation trade-offs either at a local scale (e.g. in hard real-times systems with physical effects) or at a global scale (e.g. in telecommunication systems).

The distribution and pervasiveness of computing must be taken into account when designing and deploying infrastructures, when engineering software, when collecting and processing data, and when offering services to a broad audience.

### Secure/Trustworthy/Sustainable Computing

As digital systems now support nearly all critical processes of modern societies, their non-functional characteristics have a direct impact on the lives of millions of citizens. These characteristics include security, privacy, trust, reliability and availability, resource optimisation and efficiency, and cost and performance. Maintaining nominal characteristics or acceptable degradation in case of changes in the environment or the systems themselves, or in case of malicious actions, is usually very difficult to achieve in most situations, and especially where economic sustainability is also a requirement.

Implementing these characteristics requires dedicated mechanisms in the infrastructure, in the software design and development processes, or in the way data is stored and processed. The contractual aspects of services have to include service level agreements applying to the different functions expected by users.

### Smart/Cognitive Computing

While very basic at low-level, computing systems can achieve very rich behaviours when complex algorithms, availability of pertinent data, and awareness of context combine to provide insights and to perform actions that would not have been possible without the support of digital artefacts. Tackling unprecedented levels of complexity, implementing reaction loops of different spatial and temporal span, and performing simulations and optimisations with great precision is now possible and often necessary to manage the digital substrate of society.

Domains like artificial intelligence, machine learning, evolutionary or bio-mimetic mechanisms, and high-level intention-based human-machine dialogue need to be applied at a larger scale than ever. They can be applied to reinforce the infrastructures, they lead to new ways to consider and create software, they enable new insights from data and they can lead to novel paradigms in delivering and trading services and in extracting value from the interaction between actors (humans, organisations, digital artefacts).

### Challenges

In the rest of this chapter of the SRIA, we analyse the various research and innovation challenges in the scope of NESSI according to the Dimensions and Characteristics defined above, as set out in Table 3.1 below.

| Characteristics / Dimensions | Physical Embedded Distributed Computing | Secure Trustworthy Sustainable Computing | Smart Cognitive Computing |
|---|---|---|---|
| A - Digital Infrastructures | Computing Continuum | Dynamic policies and behaviours | Self-adaptable infrastructures |
| B - Software Technologies | Full-stack software | Programming policies | Self-aware programs |
| C - Data | Data everywhere | Built-in protection of assets and actors | Knowledge and insight for decisions |
| D - Services | Software defined sectors | Regulation in societies of services | Digital business agents |

Table 3.1: Challenges by Dimensions and Characteristics

## A.1 Computing Continuum

**Vision: a consistent infrastructure that is able to integrate objects in the cloud world, through pervasive networks and adequate middleware. The computing continuum mixes centralized and decentralized CPS, cloud computing systems, fog computing, edge computing, and support for computing of smart objects (or at least integration of objects) in cloud applications using API and protocols.**

**Challenges**

The world is witnessing the emergence of new distributed computational models. While the 5G community is opening a new era for telecommunications, citizens are creating their own communities for sharing networks and computational resources: computers will be interconnected at very high speed, and latency will be dramatically reduced, no matter where the terminal and the servers are. From a business perspective it could be said that the barriers among service providers, prosumers and consumers are going to be lowered, making possible interoperability models that were not even thinkable before. Workflows will cross traditional business boundaries, connecting partners in a fluid and open computational environment. It could be said that the computational convergence will be not only horizontal, that is to say among business entities, but also vertical, from consumers toward business and the other way round.

So-called Jungle computing is another demonstration of the ongoing process toward pervasiveness and fluidity of computation. This new computational model identifies the combination of different type of computational resources working together in a coordinated manner. It is a way of letting different computational resources work together: graphical processing units, CPUs of different types, and different types of data sources. Jungle computing is thus another demonstration of the computational fluidity envisaged in a more and more connected world.

In this ‘liquid’ computational world, the mesh-up of services opens new possibilities for collaboration among different actors and for the creation of new value chains. Vertically, from prosumers toward ‘traditional’ service providers, these coordinate different type of services, different delivery models, SLAs, security profiles, etc., in a single, uniform provisioning model. These new service provisioning chains may be opportunistic, created on-the-fly using new and more and more intelligent brokerage techniques (e.g. agents, goal-oriented mesh-up algorithms, etc.), or more structured, based among the other things, on secure mechanisms for accounting the contributions of all the actors in the value chains and secure, trustable mechanisms for integrating. It is clear that, orthogonal to the problems related to the integration of services, a deep review of the methodological approach is needed.

We expect to see more and more ‘intelligence’, to capture user intention (i.e. abstraction of program code), and to deal with data and business goal abstractions (e.g. cost, data locality, performance, security, privacy, environmental friendliness, etc.). This new model, once mature, could subsume all other computing models, providing a revolutionary, comprehensive answer to still unresolved business and societal problems.

## A.2 Dynamic policies

**Vision: non-functional characteristics are addressed in the infrastructure, masking complexity, failures and risks to provide a sound foundation on which to base services.**

### Challenges

When developing components or solutions on top of a software architecture, the main focus is developing the core functionalities that are required to provide the expected service. Modern agile methodologies emphasize this focus when the target is to design a Minimal Viable Product. A side effect is that the other characteristics of a system may become less important, with trade-offs on the level of quality, security or performances that may appear to have lower priority than the visible functions of the system. One option is to be able to embed in the infrastructure the support for some non-functional aspects of the application, facilitating the job of application and service developers, thus allowing them to concentrate on the core purpose of their application. For example, cloud management must be based on generalized policy-based orchestrations that free applications from explicitly programming elasticity, recovery, reconfigurations, etc.

Embedding advanced capabilities in software infrastructure first requires implementing the software that will be able to support these functionalities, designing and deploying components and solutions that will complement the basic functions of the infrastructure and provide add-value behaviours and guarantee. One challenge is to tightly integrate the new functions with the existing ones, while preserving the consistency and quality of the global architectures. For some aspects, standard solutions exist that implement some level of technology (for example cryptography, PKI, privacy enhancing technologies): they have to be woven into the fabric of the infrastructure without introducing weaknesses or degrading their level of performance. For example, systematic encryption of data (blind storage) or interactions (APIs) must be supported by the ad hoc support of the key management in the infrastructure itself.

For applications to benefit from functions embedded in the infrastructure, using these functions must be easy for service designer and developers, without slowing down the pace of development of applications or increasing noticeably the associated costs. Automation and tooling can help in reaching this goal, by embedding complex behaviour and decision making either in the development tools or in the runtime of the infrastructure. A key point is that the components and tools of the infrastructure must exhibit the same dynamicity as the development and deployment life cycle of modern digital services, for example by using high-level policies that describe some domain rules that need to be verified constantly, in case of changes in the infrastructure, in the application or in the external context. Rule-based policies can be used to enforce security, verify privacy assertion, make decisions on performances or optimization, control the dispatching of resource, or consistently manage other aspects of infrastructures and services.

In large scale or complex systems, it is still a conceptual and technical challenge to identify the key non-functional characteristics, to automate their support, to control them using domain based explicit policies and to enforce them simultaneously in a consistent way.

## A.3 Self-adaptable infrastructures

**Vision: infrastructures can adapt to change, and in addition they are "aware" of these changes and can continuously and autonomously apply sophisticated heuristics to adjust their level of performance.**

**Challenges**

The increasing scale and variability of compute workloads (e.g. social, mobile, big data, cloud, IoT, Enterprise IT, HPC, etc.) is driving significant changes to the manageability of the underlying computing infrastructure systems. Such infrastructures have been morphing from being managed as discrete physical resources into flexible pools of reconfigurable resources (e.g. Software Defined Environments/Infrastructure), enabling infrastructure systems to scale up or down based on demand or environmental conditions. However, the physical manageability of infrastructure systems is becoming ever more complex due to such factors as distribution and disaggregation, infrastructure hyper-scaling, and infrastructure heterogeneity. For example, some mission-critical large-scale and distributed IT systems are reaching manageability saturation. There a growing need for a smart management system for provisioning and running infrastructure resources that are efficient, scalable and cost-efficient, as well as ensuring secure, reliable, and resilient services.

The characteristics of autonomous systems applied to infrastructure management have the potential to answer this need. An autonomous system can adapt to unpredictable changes, shielding complexity from operators and users whilst still being under their control. To reach this goal, such a system must rely on self-management properties including self-configuration, self-optimising, and self-healing.

An Autonomous Management System (AMS) for computing infrastructure is different from current cloud infrastructure management. Many applications and platforms, although used as services (SaaS/PaaS) directly from the cloud infrastructure, have not been designed to be dynamically scalable. An AMS could be viewed an extension or the next generation of Cloud. An AMS would be able to self-manage, given high-level objectives (e.g. dynamic scalability, infrastructure interoperability, massive parallel computing, etc.), utilising adaptive technologies.

For practical reasons, infrastructure management needs to be as autonomous as possible i.e. self-managing, self-optimising, and self-healing. An AMS would seamlessly manage infrastructure from within the datacentre, across the network and out to the edge. An autonomous system would require extensible, machine-readable, future-proof, trustworthy standards, machine learning techniques, APIs and protocols for

- Infrastructure Landscaping: express or identify currently available, recently arriving and previously unknown resource capabilities across a heterogeneous infrastructure landscape, e.g. self-discovery, registration, advertisement, etc.

- Infrastructure Workloads: expose the characteristics of various workloads and their affinity for specific infrastructure resource capabilities.
- Infrastructure Acumen: optimise infrastructure performance (and therefore the services that reply on it) by allocating resources (hardware and software) in the right amount, at the right location and at the right time, against available infrastructure resource capabilities, to deliver expected quality of service with acceptable operation costs. Furthermore, an AMS should have the ability to self-evolve through continuous learning, advancing towards an optimized model.
- Infrastructure Hyper-scale: to support hyper-scale deployments, the complete life-cycle should be supported to minimise (if not remove) human interaction, and optimise performance and efficiency, e.g. automatically determining resource setup configurations, fine-tuning over time, and release or retirement.

## B.1 Full-stack software

**Vision: designing useful applications and services will be facilitated by the capacity to both programme elementary behaviours (within objects, for human interactions) and to enable deployment, execution and management at a global scale.**

### Challenges

In a world where IoT will connect numerous cyber physical systems together and to remote services, developing consistent and efficient applications will be a challenge. Each component may have specific requirements on the tools, methods, languages used to design, develop and test the software. Embedded systems will require high levels of optimisation to meet the tight requirements on autonomy, responsiveness (including hard real time), resilience, and adaptability to local changes in the local context. These embedded systems will be managed/queried by services running in the cloud, which may be directly used by humans or integrated in more complex scenarios involving multi-domain services. To cope with the dynamicity, scalability and adaptability required by such use cases the services will be implemented using distributed frameworks and patterns, where units of computation and storage will be designed to facilitate integration, deployment and management. Micro-services (a form of service oriented architecture), agent-oriented systems, new programming models and languages adapted to a cloud computing infrastructure, and distributed middleware for resilient communication and storage are some of the solutions that may be used, along with continuous integration and deployment and advanced monitoring and administration of complex systems.

The term ‘full-stack’ comes from the deployment of mobile and web applications where it is now possible to develop all parts of the applications using common frameworks and libraries, and in some case the same languages. Software engineers can thus have a global view of the application, improving the ability to rapidly evolve and improve the global quality (e.g. diminution of interfaces and impedance mismatch situations). It is a challenge to extend this full-stack approach to systems with embedded sub-systems (captors, actuators), distributed intelligence, and information and services offered through multiple channels (mobile devices, web, API). This requires models and middleware for the composition and evolution of both CPS and micro-services, high level orchestration (including migration of process and data), and tools and methodologies to design, simulate and evaluate the software architecture and its possible behaviours. In some case (for example, very dynamic systems with a large number of interconnected CPS), multiple levels of tool-supported abstractions may be required to capture both the local reactive behaviours and the emerging properties and process of the global system.

## B.2 Programming policies / Policies by design

**Vision: non-functional characteristic are natively programmed in the system, using appropriate formalisms and models, at an abstraction level (policies, intentions) appropriate to human governance of legal aspects and social responsibility.**

**Challenges**

During the design and development phase of software-based systems non-functional as well as functional requirements should be considered from the very beginning. Non-functional requirements include e.g. security and privacy, availability, performance, reliability and response time. Such non-functional requirements are often expressed as policies or assertions that need to be verified. The code generated should meet these requirements and automatically conduct checks (during runtime) so that the system always performs as specified. Explicit definition of objectives, metrics and evaluation criteria is mandatory to assess the non-functional guarantees provided by different techniques.

Dynamic security and privacy mechanisms guarantee design and run-time assurance and support to improve the analysis and treatment of security and privacy risks. These could also be manifested in certification of the system. Novel security and privacy lifecycle and process models are required, as agile development does not fit the sequential nature, or the requirements for extensive documentation, of traditional security engineering processes. Based on empirical evidence the suitability of traditional security engineering processes in agile development settings can be proven (see the NESSI White Paper on Security and Privacy). Information about the integrated mechanisms and features should be clearly communicated to users of systems, so that they are aware of the privacy consequences of their online activities, and so that systems offer implicit support to users.

Similarly for other non-functional characteristics, high-level goals related to business and regulation have to be expressed as policies, to be implemented by design and systematically verified during the whole lifecycle of systems, using appropriate tools to program the impact of these policies and to assert their enforcement at run time (management systems, assertions and observations, data analytics…).

## B.3 Self-aware programs / World-aware programs

**Vision: programming self-adaptation through self-awareness and constant adaptation to context departs from imperative programming and uses new tools and algorithms, able to continuously co-evolve with the system they control.**

**Challenges**

In the era of cognitive computing and artificial intelligence self-aware programs will be built using novel machine learning and deep learning techniques which have significantly different characteristics compared to imperative programming tools and algorithms. By their nature machine learning techniques are not deterministic and can generate different results each time they are invoked. This creates special difficulty when developing, debugging and testing software systems that adapt themselves over time. In addition, the behaviour of self-aware programs depend on the data. The quality of data and its cleanliness will affect the quality and performance characteristics of adaptive software systems.

An additional aspect of self-aware software systems is that they often use special hardware acceleration for deep learning algorithms, in order to deliver the required precision and response time. In order to debug and test an adaptive system, the hardware acceleration subsystem has to be transparently supported by the SDK and the related debugging/testing interface. The availability of, and support for, hardware acceleration from the beginning of the software development process is crucial for the success of developing a new generation of computationally intensive self-aware programs.

Another aspect of self-aware programs is that they usually consist of a collection of communicating software components. For example, a cognitive system to assist a car driver will have a component that processes images received by the sensors, another component that processes voice commands, yet another component that uses navigation maps, and so on. Decisions made by the self-aware adaptive system depend on the inputs and on communication among all the components involved. This inherently distributed structure of the self-aware systems leads to additional difficulties in development, debugging and testing.

Fully-automated software for self-aware systems will evolve and morph during run-time due to learning and cognition, without the explicit control of software engineers. Whilst in traditional programming the code is logical, deterministic and examinable, with the rise of machine learning techniques (particularly deep learning over neutral networks), the ‘code’ may be largely opaque and inscrutable - in other words, a black box. Even simple algorithms can create emergent behaviour when they are intertwined or when feedback loops are introduced, resulting in ‘ghosts in the machine’. This raises questions about the overall behaviour of the system, and how to intervene in a fully automated system of systems in case something goes wrong. Techniques to ensure that such software will always work within safe operational boundaries are required. Current verification and validation techniques work during design time or exploit pre-defined strategies to anticipate how systems may adapt and evolve during run-time. Due to learning and cognition, such anticipation of run-time behaviour is no longer possible, and a different approach towards quality assurance is required. Advances in learning techniques (such as deep learning) combined with more powerful means to deploy verification tools (such as verification in the cloud) will provide the required means to address this challenge.

For several years NESSI has been advocating that software engineering techniques need to adapt to the new world of self-aware systems. Investment in research and innovation in this domain is essential.

## C.1 Data everywhere

**Vision: as digital processing invades objects and human processes, the data generated and used is distributed on a very large scale, requiring ad-hoc heuristics to manage information and extract knowledge and value at all levels.**

**Challenges**

In distributed information systems encompassing both connected devices and software solutions, data is used, generated and stored at different levels. Network capacity allows for the transparent distribution of data, but in many cases, for efficiency, performance or security reasons, smart patterns have to be used to implement data storage and processing, to prevent irrelevant information being transmitted. In the case of large-scale integration of numerous CPS systems and objects, handling the massive data streaming from the edge of the system requires real-time data collection, analysis and processing, occurring as close as possible to the source of data to optimize the global flow of information. Among the functions that may have to be performed are anonymization, statistical clustering and filtering, pattern matching, learning, prediction, etc.

As data enters the information system, additional concerns need to be addressed. For example, in systems where functions are distributed in a service-oriented architecture or a micro-services architecture, data must be explicitly part of the patterns that bring availability, resilience and scalability. Data management in cloud computing systems introduces new levels of requirement on the database systems, which must cope with new transactional profiles and new levels of replication, thus challenging the traditional solution in favour of new contenders (the so-called NoSql group).

In such systems, a piece of information can be at the same time part of a local control loop of an embedded system, of a visualisation application for a human operator, of an automated decision support system to optimize a process, and of a global analysis of technical or market trends at an enterprise level. Frameworks to describe such use cases have to be created, to facilitate the design of this kind of complex data management scenario in conjunction with the software development of the system.

## C.2 Built-in protection of assets and actors

**Vision: as actors (governments, organisations and individuals) measure the importance of data and the associated risks, they need to be assured that adequate mechanisms for protection are in place, through proof, auditability, certification and traceability.**

### Challenges

(Big) Data is commonly recognized as a fundamentally game changing set of technologies across many industrial sectors and societal solutions. A major challenge for Big Data is to overcome the fragmentation of the market and to achieve cross-fertilization between many established silos. On the other hand, actors are increasingly making use of online applications and services, unaware of the related data consumption. This increases the number of information items (‘attributes’) that are collected or processed, heedless of whether these items are already personally identifiable information. Thus, adequate mechanisms that prevent re-identification of actors while providing proven protection are required. Anonymization techniques must be found which can only be reverted by the data controller and at the same time do not make the resulting data unusable for performing the analytics. Ideally, the data provisioning process will already be controlled to minimize the collection and the storage of data (purpose-driven data minimization and purpose limitation). Taking performance into account, efficient new security and privacy mechanisms keeping the data layer thin are required.

New trust and policy models and technologies to protect data throughout its lifecycle spanning capture, transmission, storage (potentially redundant and remote), and destruction need to be in place. This is even more relevant in cloud environments where data is distributed over several information processing systems. Data should be accompanied by security and privacy policies that must be obeyed even when the data are passed on to a different processor.

Data analytics can be applied to automatically detect sensitive information in unstructured data, e.g. detect sensitive text in emails, as well as to detect privacy violations. It can be used to support the actor with decisions regarding which data is sensitive and to what degree it is sensitive. In particular, it can play the role of an 'early warning system' to detect whether data about a user which is distributed over several databases could be misused to reconstruct personally identifiable information.

Finally, as additional challenge privacy-enabled applications must be certified or labelled with respect to existing regulations and legal requirements.

## C.3 Knowledge and insight for decisions

**Vision: as knowledge and insight of all sorts can be derived from data, the technical aspects need to be aligned with the intentions of the systems, taking into account the individuals and organisations involved, and the possible impacts on society.**

### Challenges

In the new generation of cognitive systems the data volume is exploding, the type of data is taking on increasing forms, and the demand for speed at which we need to understand this data is accelerating. We need computers that can assess this flood of data so that we can mine the most value from it, and we need complex structures for capturing information and converting it into knowledge for enabling decision making.

Ninety percent of all the data in the world was produced in the last two years. Data volume is ever-growing as we interconnect and instrument more of our world. Eighty percent of all data in the world is unstructured, which includes text such as literature, reports, articles, research papers, emails, blogs, tweets, forums, chats, and messages. With all this data, we need better ways of understanding it to find knowledge that can be applied to solving the most pressing and important issues. We need to be able to make sense of what is being said, and to use that information to answer questions, gain insight, and drive better, more informed decisions. However, reading this information and understanding it with the same success that people can is hard for traditional computing systems. We need a new class of computing that is capable of understanding the subtleties, idiosyncrasies, idioms and nuance of human language.

Cognitive systems mimic how humans reason and process information. Unlike traditional computers that are programmed to calculate rapidly and perform deterministic tasks, cognitive systems analyse information and then draw insights from that analysis by using probabilistic analytics. They learn from their interactions with data, in effect continuously reprogramming themselves. These cognitive systems require machine-understandable representations, which are needed to allow automated analysis and reasoning for making decisions. For example, healthcare organizations can use cognitive systems to analyse all available data, especially textual information and medical images, to improve patient outcomes while making processes more efficient. Financial companies can use this technology to analyse vast amounts of unstructured data to improve credit decisions, investment analysis and risk management.

Knowledge needs to be captured in special model-based representations which will allow it to evolve over time by applying machine learning to derive new knowledge insights. Methods of correlation and analysis must be developed to maintain accuracy at acceptable levels and to provide meaningful and useful information to human operators. Such knowledge representation can be domain specific, e.g. an ontology used for knowledge representation in healthcare is different from one used in the finance industry. In addition, when reaching conclusions or advising human operators, the knowledge representation system needs to be able to back-track and show to humans why a certain decision has been made and how the machine learning system reached this conclusion. As machine learning assumes responsibility for more and more of our daily digital tasks and decisions, they could become potentially elusive and ungovernable. New forms are needed to 'visualise' and explain what is going on 'under the hood' of a machine/deep-learning system. This would not only assist with transparency and governability, but also how to more efficiently optimise such techniques for more predictable outcomes. For example, an evidence-based strategy is essential to convince medical doctors that a diagnosis proposed by computer is based on solid arguments and known facts from the medical literature.

Knowledge representation for advanced cognitive systems requires substantial research and innovation. This is essential if we want to see cognitive systems deployed in the mainstream of advanced computing.

## D.1 Software defined sectors

**Vision: digitalization spreads to all corners of society, changing value chains and business models, accelerating changes and migrations. Sectors must implement a sustainable urbanism for their digital services (pure digital or virtualization of physical ones) to facilitate cooperation and integration. This generalizes the role of digital systems in the R&I agendas of all sectors.**

### Challenges

As digitalization is profoundly transforming the way enterprises and administrations are organizing their internal processes and their relation with third parties and the way users expect to interact with products and services providers, numerous vertical sectors are going through a mutation phase that requires a major change that will put IT systems at the core of the organisation. It allows major progress in the way relationships with customers and partners are managed, in the way production and support processes are governed and optimized, and in the tools used for decision support at local or enterprise levels.

In each sector, a global architecture may emerge, based on the technical or market specificities of the domain (the way IT is used inside the products or inside the production process) and on generic frameworks that may be common to several sectors (customer relationship management, ERP, Business Process Modelling, web presence, social networks presence, integration with partner, API exposure, etc.). The architecture should be based on common concepts and ontologies to ease interactions between actors, on common models (for API or data) to help interconnect inside companies or with third parties or help automate some production or support processes, and on generic frameworks to reduce costs by leveraging cross domains solutions (on middleware, on tools, on data management, etc). As lots of domains will use CPS, cloud computing and data management in a similar way, a challenge will be to avoid the fragmentation of solutions for each sector and to promote research and innovation that applies to multiple sectors. Another challenge is to provide frameworks and solutions that can be adapted in different contexts (countries, organisations, regulations) without major software development or product evolution, and that natively support domain and services architectures, data and API valorisations and product line management. These frameworks can be integrated to build powerful environments to rapidly design and deploy applications and services: such environments can be offered in the form of Platform as a Service (PaaS), and possibly enriched with additional middleware functions as extended PaaS (xPaaS) to solve generic or domain-specific problems.

One additional challenge is to provide solutions to help the smaller actors in a domain who may not have the in-house skills to take charge of the transition: software and tools as a service, cloud hosting of applications and processes, and externalisation of data management and analysis could be used and promoted after a collaborative qualification phase to ensure that all specificities of each domain can be supported and that technical ecosystems can be built to augment the synergies.

## D.2 Regulations in societies of services

**Vision: digital services thrive in interwoven 'coopetiting' ecosystems, supported by a large variety of actors. Services can be dynamically aggregated, integrated and replaced, as long as fundamental mechanisms of interaction are preserved.**

### Challenges

As the service-based and knowledge-based economy gains importance worldwide, it creates ecosystems - usually supported by digital platforms - that organize the exchange of services and data; ease the interactions between actors, from the discovery phase to the contractual phase; accelerate innovation, by providing a global framework and some commons rules and solutions; and provide economies of scale for all participants. Examples of such ecosystems can be seen in cloud computing, mobile operating systems and devices, IoT, around major web actors through their API programs, and in open source.

As a growing number of actors rely on the support of such ecosystems for their processes and for integration in their own products and services, technical integration through data, API or software components is no longer sufficient. The interaction of actors through services in these ecosystems must take into account other aspects such as accountability, traceability, security, privacy, compliance to regulations, risk assessment and insurance, etc. The scenarios to be supported range from simple hub-and-spoke ecosystems to very complex hierarchical or horizontal value-chains, and include multi-tenant platforms. All traditional aspects of 'bricks and mortar' services will need to be taken account of in these digital ecosystems, with additional rigor and modelling; this will require the extension of the metrics and rules used currently in SLAs, as well as some solutions and components to automate the discovery, negotiation, supervision and litigation phases. This may lead to new forms of organisations or market places, such as distributed autonomous organisations or communities with decentralized governance. Providing some levels of access or control for the authorities will create new challenges, especially in the case where fully distributed frameworks such as blockchains are used to support transactions or used as ledgers. These new technologies will radically change the way local and global interactions and transactions are managed, potentially transforming the enterprises, the economy and even democracy.

## D.3 Digital business agents

**Vision: services have a personality of their own (behaviour, responsibility, impact on economy). Important functions and businesses are performed by autonomous agents, with loose supervision from humans, and new business models emerge.**

### Challenges

With the advent of autonomous decision-making algorithms (automated trading, and high-level control of large-scale systems like networks or utilities), the progress in knowledge management and predictive systems (machine learning, advanced analytics), and the use of artificial personalities to interact with humans (chatbots, avatars), a growing number of activities can now be performed by digital agents. A first challenge is to help organisations develop the agents that will improve their efficiency and competitiveness, covering the technical aspects (knowledge representation, decision support, control loop) and the integration with the existing systems, and also the business and legal aspects (contracts, billing). A second challenge is to improve the way these agents are perceived and understood by the human employees that need to interact with them or control them, by users and customers that may rely on them, and by authorities that need to take them into account in regulations and litigations. The technologies used to build smart agents, the amount and nature of date they process, the algorithms used and the intention they support should be exposed and formalized, to help improve their acceptance on a large scale.

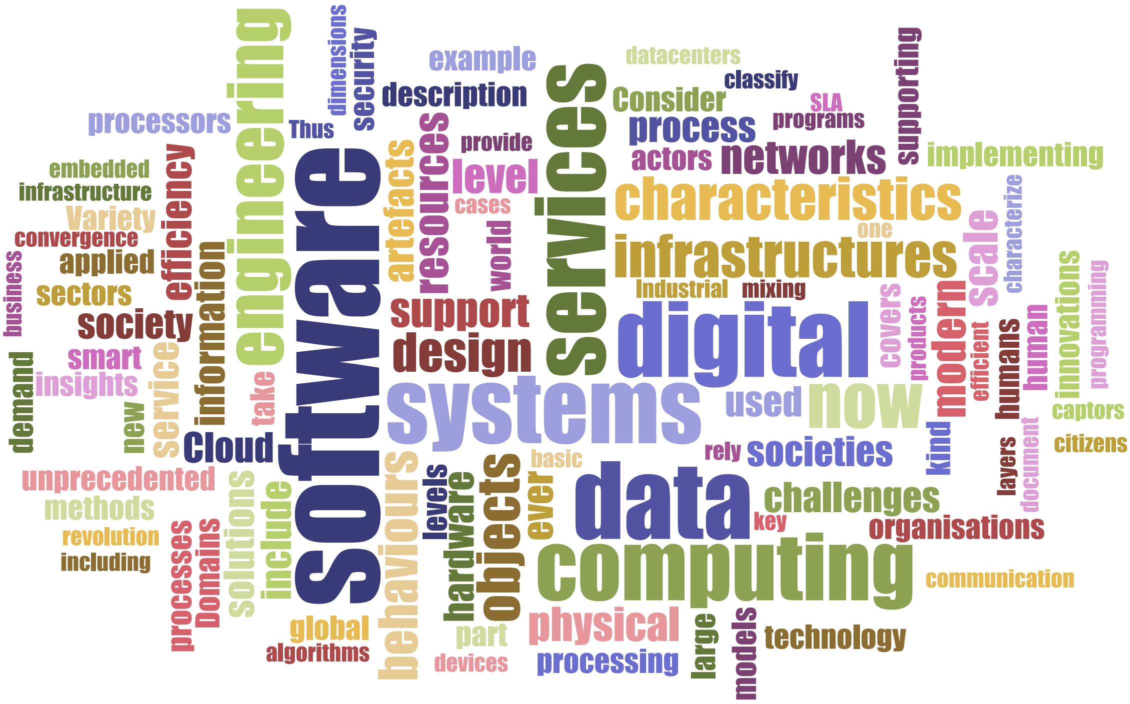
New type of organisations will emerge that will directly benefit from software platforms for rendering their services or managing their activities, such as distributed autonomous organisations based on new trust models such as blockchains. Human and agent decisions and actions will be natively supported by digital artefacts, mixing advanced data management and smart global algorithms.

An additional challenge is to monitor and evaluate the impact of digital business agents on the economy and society, through multidisciplinary studies mixing IT, economics and sociology, through simulation and identification of emerging behaviours at micro-economic or macro-economic scales, and through the study of the transformation of value chains that go through a digitalisation process. New scientific and technical models may be required to address difficult subjects like cross-sector networking effects, virality of information and usages, complex market places, medium term predictions, etc.

New multidisciplinary technical domains and research areas may emerge to address all these challenges, such as large scale software simulation of economy or society, software-aided politics, education or economy, regulation on rights and constraints for autonomous systems, formal modelling and checking of laws, and automated law enforcement for digital agents.

## Summary

The word cloud below is composed of keywords from the above sections, and summarises the research and innovation challenges in the scope of NESSI.



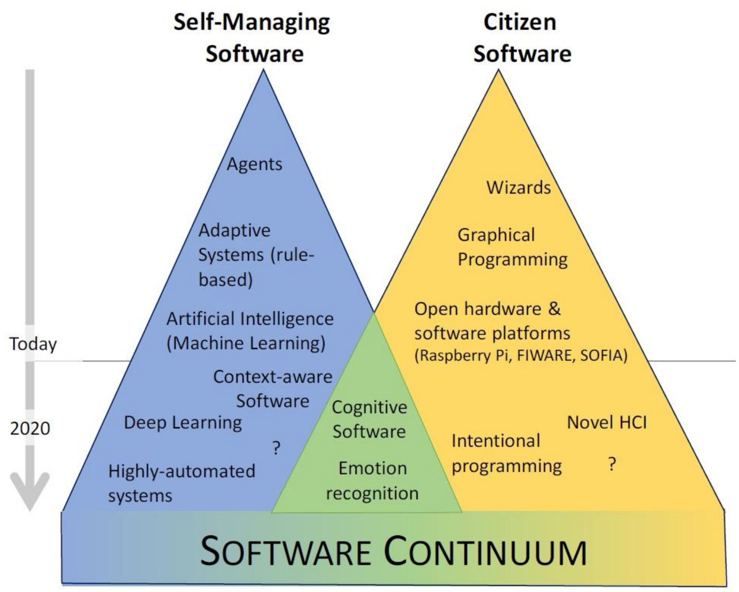
# 4. Instruments, ecosystems and dissemination

## Recognizing the importance of software

NESSI strongly believes that Software Technologies should be considered as a Key Enabling Technology (KET) as they are the foundations to achieve a full digital stack and the cross-sector interoperability underlying the EU's ambition of digital transformation of both society and industry[1].

Software technologies are the catalyzer and the accelerator for a positive feedback loop on innovations and usages. Jevons paradox states that when technological progress increases the efficiency with which a resource is used, the amount of resource necessary for any one use is reduced, but the overall consumption of the resource rises because of increasing demand - caeteris paribus. Advances and innovations in software and software engineering methods contribute to efficiency improvements in the utilization of technology and data as resources for generating services. This in turn leads to more technology and data being used, boosting the creation of more innovative services. Thus, each investment in software technologies can economically leverage a chain of cross-sector innovations and result in a family of new services.

However, the impact of software is not limited to the technical aspects of our society. It will also modify the way people interact with digital systems, organize themselves, manage their environment, and make local or global decisions. NESSI proposes a holistic view of the mega-trends at play, advocating the vision of a ‘software continuum’ which is going to provide the ‘neural system’ of the digital world, and introduce new capabilities for a sustainable connected humanity. Hence, this will raise the stakes far beyond the mere technical quality of systems to the ethical and political aspects of society. Targeting a software continuum as the future of digital services creates new challenges as to how services are supported by innovations in digital infrastructures, software technologies, and data. Thus, software innovations fuel innovations in all sectors, and transform human behaviour through education and usage.




[1] Digital Single Market, and Digitizing European Industry.

Figure 4.1. Software Continuum linking Systems and People

Targeting a Software Continuum is of paramount importance when building smart cognitive systems or designing digital services. Software is the glue that can consistently integrate all aspects: domain specific characteristics and constraints, human expectations and behaviours, richness of data, and complexity of algorithms. In an innovation process, all such aspects have to be taken into account to accelerate the uptake and impact of research and innovation. It also require an interdisciplinary approach, integrating technological innovations and adaptation into human ways of thinking and behaviours.

### Organizing Research and Innovation on software and digital services

NESSI has consistently advocated that Software Technologies are everywhere[2]. However, there are two distinct and important perspectives of contemplating software in such a context. The first is to look at software technology, like other technologies, as a means of developing and producing challenging solutions (e.g., products, systems, services). The second is to look at software technology as having challenges on its own to meet new demands, which this SRIA has explored in detail in Chapter 3, above.

Therefore, NESSI strongly recommends that the European strategy for research and innovation explicitly addresses the evolution of Software Technologies in its own right (*yesterday's state of the art cannot produce tomorrow's software solutions to meet future demands of businesses, society and individuals*), and moreover that R&I actions in domains heavily relying on software - such as Cloud, Cyber Physical System, Internet of Things, High Performance Computing, 5G Networks, Data, Digital Security, and Next Generation Internet - all include dedicated items on the necessary advances of Software Technologies to meet the digital challenges of their domain.

Similarly, in several of the Societal Challenges we encounter the need to advance Software Technologies to tackle application- and sector-specific challenges, for example software safety certification within Health, software security and privacy by design within Secure Societies, software real-time responses within Mobility, etc.

NESSI therefore envisions synchronisation across DGs, Directorates and Units as a ***virtual action on Software Technologies***. Such an action is necessary to maintain and advance Europe's position in building skills and utilizing Software Technologies as a KET in all sectors. This action should be explicitly expressed, with dedicated budget allocations, e.g., under each specific action where challenges involving Software Technologies appear. It should maximize the impact of software innovations on the digital transformation still necessary to meet all the challenges facing Europe in the next decade.

### Improving synergies in the European software ecosystem

NESSI will contribute to the software and digital R&I ecosystems at the European level, in a cross-cutting and open way. Further cooperation with other actors in the European Research and Innovation ecosystems are ongoing; such as several EC Units, PPPs, other ETPs, other initiatives (ECI, procurement, DSM), taking into account the transversal nature of NESSI's domains. Given their importance for producing and disseminating results of research and innovation, open source, standardisation and education must also be considered.

Open source can be seen as a way for universities and research labs to improve their impact and their transfer rate from research to products and as a way to help SMEs develop their

[2] e.g. NESSI Session at ICT 2015 entitled "Everything as a Service", www.nessi.eu/?Page=ict 2015.

digital maturity and independence (both in the solutions used and the skills acquired). At a global level, commoditisation of basic components through open source cannot be escaped, requiring a continuous flow of innovation to sustain the digital economy. The current dynamics of software and digital industry prove that a variety of business models can and should coexist, inducing adequate investments, growth opportunities and market/society acceptance. Open source can thus play an important role in the digital innovation process.

For digital sectors, standardisation can also play an important role to help disseminate innovation and create value. The SDO (Standard Defining Organisations) landscape in ICT is now very complex and dynamic, some traditional organisations getting officially involved in open source, some closed “clubs” being able to act as game changers. Given the pace of innovation of the digital technologies, new way of working between the research and innovations process and the standardisation/disseminations processes should be found: direct involvement of SDO in calls definition, participation to consortium, common brainstorming and road mapping activities.

In the longer term, for the next Work Programme (FP9 – 2021-2028) we need to re-think Software Technologies, due to the digital transformation we now encounter on a global scale. Europe should be in the pole position fully taking on the digital transformation and being a forerunner.

Before going into FP9 challenges, we shortly elaborate the ***three*** most fundamental and cross discipline challenges that Europe has to face over the next decade:

- The digital skill factor affects how individuals are successful in the work market and how companies compete in the world market.
- Software programming must be part of curricula in primary and secondary schools and be applied to other scientific disciplines (modelling simulation, analytics, virtual experiments…).
- Software Technologies needs to become part of cross discipline courses in higher education to foster the digitalisation of all domains.

## Shaping the digital future

NESSI's vison on Software Technologies in FP9 is that it becomes the "pen and paper" of future professions. This vision can best be met with the following objectives for future Software, Services and Data as key enablers:

- *We go from Cloud Computing* **to** *Large Scale Infrastructures offering Distributed Exabyte Computing* - the edge, fog, liquid computing paradigm will show their capabilities in domains such as health, intelligent transportation systems, manufacturing, energy, social media, etc.
- *We go from a dedicated delivery software/service model* **to** *We provide you with a service, just tell us what you need* - as individuals we will mostly operate on the net, but still be "physically" around, virtually stimulated as a "F2F neighboring asset" and helped by digital assistants leveraging smart objects
- *We go from the Wikipedia model* **to** *the Cognitive-pedia model* - We want answers, not only search results. Decisions and guides based on advanced software technologies, relevant data and smart algorithmic results, ranked according to our preferences and context, cracking the complex problems and challenges facing societies in Europe and all over the world.

# Bibliography


**NESSI documents**

Strategic Research and Innovation Agenda - April 2013
http://www.nessi-europe.com/Files/Private/NESSI_SRIA_Final.pdf

Vision and Manifesto - June 2013
http://www.nessi-europe.com/Files/Private/NESSI_Manifesto_2013.pdf

Software Engineering white paper - July 2014
http://www.nessi-europe.com/Files/Private/NESSI_SE_WhitePaper-FINAL.pdf

Input for the consultation on Cloud Computing – September 2014
http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=6911

Response to the consultation on Cloud Computing – October 2014
https://ec.europa.eu/digital-single-market/events/cf/cloud-computing-software-engineering/item-display.cfm?id=14185

Response to the consultation on Software Engineering and Open Source – October 2014
https://ec.europa.eu/digital-single-market/events/cf/cloud-computing-software-engineering/item-display.cfm?id=14058

Big Data Value SRIA v1 – October 2014
http://www.nessi-europe.com/Files/Private/EuropeanBigDataValuePartnership_SRIA__v099%20v4.pdf

Cyber-Physical Systems white paper – October 2015
http://www.nessi-europe.com/Files/Private/NESSI_CPS_White_Paper_issue_1.pdf

Input for Work Programme 2018-2019 – Software continuum – February 2016
http://www.nessi-europe.com/Files/Private/NESSI_Input_for_WP18_RELEASE2_to_EC.PDF

Security and Privacy white paper – March 2016
http://www.nessi-europe.com/Files/Private/NESSI_Security_Privacy_White_Paper_issue_1.pdf

Response to the consultation on Cloud Computing – October 2016
http://ec.europa.eu/newsroom/document.cfm?doc_id=19406

**Other relevant sources**

Big Data Value SRIA v3 – January 2017
http://www.bdva.eu/sites/default/files/EuropeanBigDataValuePartnership_SRIA__v3.pdf

Consultation on Cloud Computing: documents and reports – September 2016
https://ec.europa.eu/digital-single-market/en/news/consultation-cloud-computing-research-innovation-challenges-wp-2018-2020