BIG DATA CHALLENGES IN SMART MANUFACTURING INDUSTRY

A Whitepaper on Digital Europe Big Data Challenges for Smart Manufacturing Industry

Version 2020

BDV BIG DATA VALUE ASSOCIATION

FOREWORD

In light of an increasing digitalization, usage of data to generate value based on artificial intelligence and the resulting requirements for the manufacturing industries, Europe is facing unprecedented challenges. For the next years, the real and virtual world will continue to merge, and the physical value chains will be supplemented by digital value chains. Lifecycles of products will be integrated from idea, design and design for manufacturing over manufacturing up to after sales and on-site customer services.

These digitalization technologies will be the basis to optimize the entire value network within a company and across companies and thereby raise the effectiveness and efficiency of processes to a new stage of organization and management. Furthermore, these technologies are the enabler for new value propositions and additional revenue streams, which will result in new business models – enhancement of existing offerings as well as new digital offerings.

It is obvious that data sharing and resulting data usage must be a win for everybody all in the value network – the data provider, the data consumer and everybody providing services to analyze data.

Beside the technological challenges, e.g. data access (sometimes in real time), data choosing and cleansing, generation of value out of data with algorithms, performing in a controlled way actions to enhance the value creation processes up to such constraints as functional safety, cybersecurity and ethical questions we are facing also skill questions, legal aspects and cultural and strategic challenges.

Big Data and Artificial intelligence shifting the balance of power in the shortest possible time. Here we must see how we can assert and expand our position very quickly. Today Europe is pacemaker for manufacturing, tomorrow Europe must be the pacemaker for Smart Manufacturing- in industrial value generation out of data and industrial artificial intelligence. In Europe the domain knowledge is available, and Europe has established a powerful network between start-ups, small and medium sized enterprises, big companies, research institutes and the government – and not to forget our Public Private Partnerships!

BDVA has a very active and dynamic group on Smart Manufacturing Industries focused on identifying challenges and opportunities that Data and AI bring to our European Industries and proposing strategic roadmaps and guidelines. This paper reflects the work developed by this Subgroup in collaboration with other communities such as EFFRA, AIOTI and euRobotics. I believe this paper brings relevant value to support policy makers and industry overall in the uptake of the Data-driven and AI technologies.

Thomas Hahn
BDVA President

EXECUTIVE SUMMARY

Over the last few years, digital technologies have transformed the economy and society, affecting all sectors of activity and the daily lives of all Europeans. Data is at the centre of this transformation and more is to come (A European Strategy for Data, EC COM 66, February 19th 2020). The Digitising European Industry EC communication has recently paved the way towards digital transformation of every enterprise, independently of its sector, size and location in the EU. Factories of the Future will be more and more "data hungry": manufacturers are realizing the wealth of still un-exploited data generated by innovative production processes and are focusing on mastering data-centric factory floor processes. This is done in order to become proactively disruptive in their market segments, to enable transformative business models with demand-driven manufacturing, and engage, for example, in agile production processes, lot size one, and mass customization.

In that techno-business context, this edition of the SMI Whitepaper "Big Data challenges in Smart Manufacturing Industry" followed a twofold approach. On the one side, an evolutionary approach has been chosen in continuity with the 2018 edition[1]; on the other side more revolutionary aspects have been analysed as bridges towards new Research & Innovation activities in the next Multi-annual Financial Framework (MFF) of the European Commission, Horizon Europe and Digital Europe in particular.

An updated and deeper market analysis, thanks to IDC surveys and DATABENCH project insights, has been performed. The research confirms that Manufacturing is one of the leading industries for adoption of Data Technologies, which is widespread across all business functions with a prevalence in product management, maintenance and logistics, and IT and data operations. Since there is little systematic evidence about the business impacts of Big Data, a key result of the survey is the substantial business benefits achieved by the surveyed enterprises, such as, gains for manufacturing enterprises are around 7% for profit, revenues and costs reduction (mean values). These impacts are slightly higher than the same value for the overall cross-industry sample. The range of variation of KPIs improvements is much wider, with several organizations achieving revenue and profit growth higher than 10%. Moreover, these benefits are expected to increase rapidly during 2020 as Data Technologies scale up in the organization.

The 2018 technical challenges, divided in to the three Grand Scenarios of i) Smart Factory (adoption of Data Technology in the production plants); ii) Smart Product (adoption of Data Technology along the lifecycle of products) and iii) Smart Supply Chain (adoption of Data Technology integrating the various stakeholders in the value chain), have been grouped into categories and subjected to a polling / voting procedure by our SMI constituency. In the Smart Factory scenario, higher priority has been given to technical challenges regarding computational continuity machine-edge-cloud and how to conveniently distribute, depending on the final application, the computational and data storage burden among the three levels. In the Smart Product scenario, higher priority has been given to the generation of multi-stakeholder Data Spaces along the lifecycle of complex products (such as machine tools, aircrafts, ships, cars) and the need for highly distributed and secure interoperability between Internet-of-Things "Digital Twins", and Internet-of-People "Digital Personae". In the Smart Supply Chain scenario, emphasis has been given to B2B industrial data platforms allowing not just trusted and secure data exchanges in static manufacturing value chains, but also to complete end-to-end data sovereignty, regarding access and usage of the data in dynamic marketplaces scenarios. A further alignment with EFFRA *ConnectedFactories*[2] project pathways has been implemented and a plan to evolve this framework including evolutionary digital transformation journeys of the new *ConnectedFactories 2* project developed.

While the 2018 paper focussed just on EFFRA and Factories of the Future communities, this version, broadens the perimeter to surrounding research and innovation communities in order to define relevant

---

[1] http://www.bdva.eu/sites/default/files/BDVA_SMI_Discussion_Paper_Web_Version.pdf

[2] www.connectedfactories.eu

cross-disciplinary challenges in the Data for Manufacturing domain. The relationship with the IoT community has been enhanced by frequent meetings and knowledge exchange (e.g. world café), by opportunities with the AIOTI (Alliance for IoT Innovation) association, especially focussing on Standards, Security and advanced Communication facilities (such as 5G and TSN Time-Sensitive Networking). Strong links have been established with Robotics (euRobotics) and AI communities with the aim of defining a strategic research, innovation and deployment agenda for an AI PPP in the next MFF. This activity started in 2019 and it is one of the major pillars of the 2020 discussions.

A more disruptive approach aimed at identifying and analyzing new trends and technologies, which could sensibly affect a Data-driven Digital Transformation of the EU Manufacturing Industry has also been reported in this version of the paper.

Firstly, this paper looked into non-technical challenges generated by an extensive and pervasive adoption of Data Technologies in Manufacturing.

At policy, regulatory and legal level the European path to Data Technologies for the Smart Manufacturing Industry needs to be driven by compliance with EU legislation, principles and values, strongly supporting a human-centric approach. It shall be grounded in the rule of law and in the protection of fundamental rights, such as human dignity, non-discrimination, freedom of expression and privacy protection. In line with the EC's White Paper on Artificial Intelligence[3], a European approach to excellence and trust is expected to play a key role in achieving the Sustainable Development Goals and in supporting the societal wellbeing, aligning its further research and deployment actions with the EU objectives to achieve an "ecosystem of excellence" coupled with an "ecosystem of trust".

At Business Model level, the paper emphasises the analogy between the very successful ongoing service Economy for Manufacturing Industry and the new born, still-to-be-developed Data and Platform Economy. Product-Service Systems are the major revenue stream for many industrial sectors, in particular for large enterprises. However, a holistic approach to their lifecycle is still missing among SMEs. The economic interest and benefit of a data economy is still in its infancy for manufacturing and B2B in general, while has recently achieved unexpected relevance in B2C and G2C scenarios. We expect an important evolution in "Data for Manufacturing" business models in the next years, to be discussed starting from the 2020 edition of the paper.

At the skills and competencies level, the main requirement is for promoting the development of an EU competence framework and an EU observatory for skills needs assessment and anticipation related to big data and artificial intelligence adoption in smart manufacturing industry. This will include creating EU international degrees and masters, modernizing technical & vocational education and training, promoting training-on-the-job programmes on big data and AI in manufacturing. To this end, collaborative schemes based on industry-research-education-government alliances must be fostered.

Finally, new technological trends that emerged during 2019 have been analysed in the Data Technology for Manufacturing Industry perspective. In particular, the evolution of RAMI, from a static reference model to a dynamic set of Asset Administration Shells and the diffusion of Digital Twins technologies, is the new empowering vision of a common EU Industrial Data Space to be developed in the next few years. Transnational initiatives, such as GAIA-X, are also moving in this direction and they will be more closely followed in the 2020 activities of the BDVA SMI group.

We hope you will enjoy reading this edition of the Smart Manufacturing Industry paper and be interested to contribute to the next one.

---

[3] https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

Table of Contents

1 INTRODUCTION

This is the second edition of this paper about Big Data challenges in the Smart Manufacturing Industry. In the **2018 edition**, the focus was on the alignment between the BDVA Reference Model and the EFFRA Strategic Research and Innovation Agenda through analysis of their respective reference architectures and alignment of their respective technical challenges. The main outcome of the 2018 activities was the identification 56 Big Data technical challenges in the three Manufacturing Grand Scenarios of Smart Factory, Smart Product and Smart Supply Chain.

Now **in the 2020 edition**, a twofold approach was chosen: on the one side (continuity evolutionary approach), we are updating the previous content, not just of the technology but also of the legal-socio-business landscape. On the other side, H2020 is in its last period and we need to think of more disruptive and revolutionary topics to be implemented along the future Multiannual Financial Framework (MFF 2021-2027) and respectively in both Digital Europe and Horizon Europe programs. This second approach will continue in the work of the SMI group on elaborating the bridge between H2020 and Horizon / Digital Europe.

The **evolutionary approach** acts in continuity with the 2018 edition by:

- An updated Market analysis about the ever increasing relevance of BDT (Big Data Technologies) full adoption for the global competitiveness of EU Manufacturing Industry (Chapter 2);
- A revised analysis of the EFFRA Connected Factories pathways and maturity levels[4], aiming at identifying and synthesising the major gaps to be filled towards 2025 Data-driven Autonomous & Smart, Hyper-connected and Collaborative Product-Service Factories of the Future, as well as at harmonizing and prioritizing the 2018 Technical Challenges for 2021-27 implementation (Chapter 3);
- An extension of the SRIA (Strategic Innovation and Research Agendas) documents to other relevant communities and domains, such as the Industrial IoT domain inside **AIOTI**, the robotics-driven agile production inside **euRobotics association** and the innovative distributed architectures and applications inside **AI for Manufacturing** (Chapter 4).

The more **disruptive approach** aims at identifying and analyzing new trends and technologies, which could sensibly affect a Data-driven Digital Transformation of the EU Manufacturing Industry in the coming years:

- The consideration and assessment of **non-technical challenges** (regulatory and ethical, business and social), which beyond the expected maturity of BDTs could severely affect the pace and modality of the digital transformation of industrial products, processes, organizations and workspaces (Chapter 5)
- Some **highly innovative technological trends** could revolutionize our current perception of the Manufacturing Industry and its vital business processes, such as the transition from static Industrial Assets data models towards dynamic (RAMI) **Asset Administration Shells**, the great success of **Digital and Virtual Twins** and the continuous and vital need for the SMI sector to access secure **Data Sharing Spaces** where data exchange can occur in a trusted Data Sovereignty infrastructure (Chapter 6).

Topics related to AI for Manufacturing and Human-AI interaction, edge-cloud-HPC computation continuity infrastructures as well as cyber security challenges along the whole value chain will be addressed in the next edition of this paper, which will be discussed throughout 2020, mostly in the BDVA AG workshops and EBDVF2020 in Berlin.

---

[4] https://www.effra.eu/connectedfactories-project

2 MARKET AND TREND ANALYSIS

2.1 Main Trends

Factories of the future will be more and more "data hungry", according to IDC predictions[5] of main trends. Manufacturers are realizing the wealth of still un-exploited data generated by innovative production processes and are focusing on mastering data-centric factory floor processes, to become proactively disruptive in their market segments, to enable transformative business models with demand-driven manufacturing, and engage in agile production processes, lot size one, and mass customization. In practice, this means data will be at the center of and will permeate every key manufacturing process step:

- In the input phase, raw material and components arrive with embedded information about their provenance and suppliers as well as assembly instructions;
- In the production phase, the continuous interplay and information handover between workers, machines, and business applications will shape every action and make every step digital and visible in real time;
- In the output phase, many factories will be dedicated to producing data centric products - connected and smart products. These products will report their status back to the plant while in operation to "request" a spare part to be produced for a maintenance job, or to report a quality issue that can be potentially addressed on the actual product via a rework in the factory (or on the product line) via process fine-tuning.

In this context, "Data first" will gradually become the design principle of new plants and the guideline driving brownfield improvements: this will require IT leaders to step up and be involved in the plant design steps. In addition, ensuring data flows across the plant will be as essential as process execution optimization, and this will create new pressure on IT. IDC expects that by 2022, 75% of G2000 manufacturers will have established digital platforms to unify product and manufacturing process data, to deliver complex products through networked processes.

Data-driven innovation is also behind other relevant trends monitored by IDC, including the accelerating convergence of IT and OT (operational technologies, including control, automation and manufacturing systems), the incorporation of customer data to increase production flexibility (with the help of AI) and the increasing use of testbeds.

The convergence of IT and OT is more than a trend. In short, it is the main enabler of smart manufacturing: the integration of enterprise and shop applications enable data-centric computing and smart manufacturing platforms. The main benefits of convergence include the improvement of operational performance (throughput/service reliability at same or lower costs), more efficient resource sharing, better and more comprehensive security, improved product and service quality, and of course greater agility and flexibility. This convergence has been an ongoing process for several years but is now reaching the upper levels of the organization, through integrated IT/OT governance models, where investment decisions regarding control systems and execution systems are made through a shared services organization, a center of excellence (COE), or a corporate function. In addition, decision making about investment and priorities for operations is undertaken as a single unit. Within three years, 50% of European enterprises should have an integrated IT/OT governance model. In addition, IDC expects that 40% of manufacturers by 2022 will employ a cloud platform that crosses traditional IT boundaries and integrates operational technology.

---

[5] IDC FutureScape: Worldwide Smart Manufacturing 2019 Predictions https://www.idc.com/getdoc.jsp?containerId=EMEA44385118

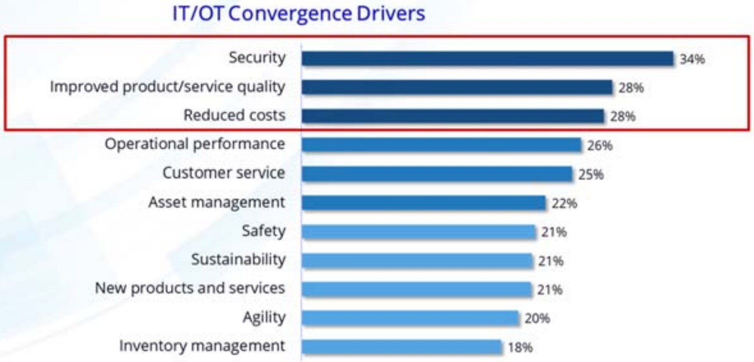


Figure 1 Europe - Main drivers of IT/OT Convergence

Source: IDC's IT/OT Convergence Survey, June 2018-
% of respondents - N = 255 (European respondents)

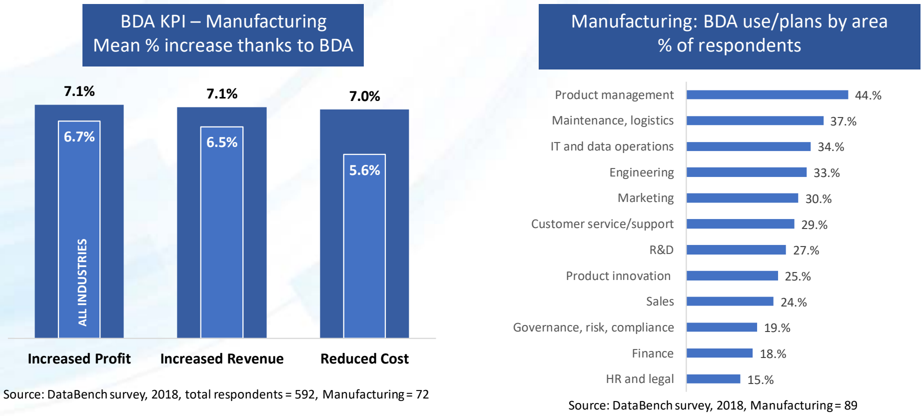
Meeting demand fluctuations and the requirements of mass customization is a constant challenge, which has led to a focus on increasing agility and responsiveness of production processes. The progress made by implementing modern MES, scheduling applications, and lastly, industrial Internet of Things (IIoT) platforms have been enormous. Now the new frontier is to seamlessly incorporate key customer data and demand signals in to factory production, leveraging data flows from connected products, customer interactions and social media analytics. The next step will be to leverage AI and automation: for example, inputs from AI-powered decision making, or flexible tools which can perform variable tasks without the need for human intervention, such as intelligent co-robots, 3D printers, and machines with re-tooling capabilities. Data centric processes also include fully interoperable and integrated processes. In order to deal with this evolving environment, manufacturers worldwide are increasingly choosing to participate in collaborative testbeds with their peers, particularly to develop/select effective interoperability standards and drive innovation on the shopfloor by testing concrete use cases without taking "leaps of faith". IDC has noticed that this experimental approach to innovation is becoming mainstream, thereby confirming the validity of the EU collaborative research support of business trials run by Large Scale Pilot projects.

2.2 KPIs of BDA business impacts in Manufacturing

The Manufacturing industry is achieving substantial business benefits from the adoption of Big Data and Analytics, as shown by the survey carried out by the DataBench project in October 2018 (Figure below), with better results than the average European industries[6]. DataBench, whose main goal is to bridge the gap between technical and business benchmarking of Big Data technologies, investigated in depth the level of adoption and the impacts of Big Data and Analytics (BDA) by European industry, through a dedicated survey of 700 European businesses in 11 EU Member States carried out in October 2018. These companies were selected after confirming their actual or planned use of BDA. The sample is representative of European industry (excluding the Public sector, Education and Construction). The research findings confirm that Manufacturing is one of the leading industries for adoption of BDA, which is widespread across all business functions with a prevalence in product management, maintenance and logistics, and IT and data operations. Since there is little systematic evidence about the business impacts of Big Data, a key result of the survey is the data on the substantial business benefits achieved by the surveyed enterprises: gains for manufacturing enterprises are around 7% for profit, revenues and cost reduction (mean values). These impacts are slightly higher than the same value for the overall sample. The range of variation of KPIs improvements is much wider, with several organizations achieving revenue and profit growth higher than 10%. Moreover, these benefits are expected to increase rapidly from 2018 as BDA scales up in organizations.

---

[6] See deliverable "Preliminary Benchmarks of Industrial Significance of Big Data" https://www.databench.eu/wp-content/uploads/2019/02/databench-d2.2-ver.1.0.pdf



Source: DataBench survey, 2018, total respondents = 592, Manufacturing = 72

Source: DataBench survey, 2018, Manufacturing = 89

Figure 2 Manufacturing BDA adoption and business impacts, DataBench survey, 2018

2.3 Debunking the Myth of Automation killing Jobs in Manufacturing

Will robots steal jobs from people? According to IDC[7], this will not happen and the most likely impact of increasing automation on factory floors will be human-machine interaction and humans working smarter.

There are two ways to potentially replace human work: deploying robots or leveraging software to automate back-office activities. At the moment, the most commonly deployed technologies are software automation technologies, used among the least specialized skills in manufacturing. As shown by the Figure below, based on IDC's *EMEA Future of Work Survey* carried out in November 2018, there is strong evidence of these technologies leading to employees working smarter, rather than to job redundancies — a well-known concern in industry. However, many jobs will be displaced/transferred to other functions, forcing companies to constantly reinvent job roles as they progress on their digital journeys. Robots for asset operations/maintenance and for remote hazardous operations show the strongest effect of this "man with machine" impact, although the deployment of these robots inevitably leads to minor job losses. Another noticeable impact of automation technology is that, in 30% of cases, it led to the creation of new jobs that didn't exist before, as workers are increasingly doing less predictable physical work, data processing, and information collection, and more decision-making tasks based on data collected in manufacturing plants.

Every industrial robot requires mechanical, electrical, and software care, which will also contribute to a shift from low/unskilled labor to higher-level opportunities over time. This will call for reskilling programs, which we are already beginning to see evidence of.

Figure 3 Impacts of automation technologies on work in Factories

Source: IDC's EMEA Future of Work Survey, November 2018

---

[7] Man with or Versus machine? Tackling Productivity in European Manufacturing (IDC #EUR145026619, May 2019)

3 DATA-DRIVEN SMART MANUFACTURING VISION TOWARDS 2025

The vision of data driven smart manufacturing has evolved from a number of research activities. Works in the Connected Factories CSA and the Digital Shop-floor Alliance led to the so called Factories of the Future 2025 Pathways (§3.1). Based on these results, a synthesis and prioritisation of the technical challenges for the Smart Manufacturing Industry can be derived that takes thoughts from the 2018 edition of this Discussion paper into concrete propositions (§3.2).

3.1 The Evolution of Connected Factories 2025 Pathways

In §3.1.1 the evolution of the three CF pathways during 2019 is presented as well as a perspective for new 2020 pathways to be developed within the new Connected Factories 2 CSA started on December 2019.

In §3.1.2 a data-driven architecture for Smart Factories is reported under the Digital Shopfloor Alliance initiative federating some Factories of the Future projects and paving the way towards Digital Manufacturing Platforms and the upcoming Horizon Europe program.

3.1.1 Evaluation of the pathways by Connected Factories CSA

ConnectedFactories CSA[8] has recently held its final Review Meeting. In this section, the latest evolution of the three Big Data Technologies adoption pathways will be reported. In December 2019 a new CSA was launched (ConnectedFactories 2) in the domain of Digital Manufacturing Platforms with the aim to co-ordinate the projects funded under DT-ICT-07 in Manufacturing. New and adapted pathways will be developed in the CF2 project and consequently discussed and reported in the 2020 edition of this paper.

Pathways to the digitalisation of manufacturing reflect how digitalisation and eventually the deployment of digital platforms can bring value within different kinds of manufacturing areas, such as factory automation, value networks or product-service development. The pathways enhance awareness among different stakeholders about the actual and future use of digital technologies in manufacturing and facilitate the migration from legacy situations towards innovative approaches.

Initially three generic pathways with a particular focus have been developed:

- The **Hyper-connected Factories pathway:** towards networked enterprises in complex, dynamic supply chains and value networks
- The **Autonomous Smart Factories pathway:** towards optimised and sustainable manufacturing including advanced human-in-the-loop workspaces
- The **Collaborative Product-Service Factories pathway:** towards data-driven product-service engineering in knowledge intensive factories

The **pathways** and in particular the associated examples (or cases) also reflect how different digitalisation approaches will or can be implemented and co-exist in a concrete business environment. The pathways developed **indicate how research and innovation projects contribute to the future deployment of digital platforms** for manufacturing and will **facilitate and stimulate the discussion and identification of more company-specific innovation strategies**.

---

[8] https://www.connectedfactories.eu/project-news/connectedfactories-csa-publishes-key-deliverables

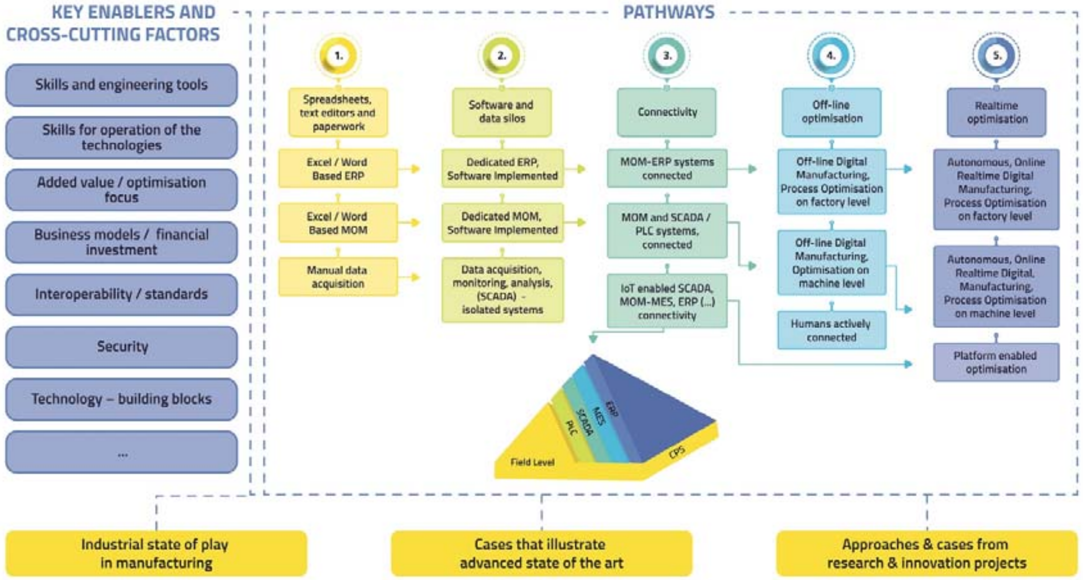


Figure 4 The pathway on the right of the picture above is the Autonomous Smart Factories pathway

There are a lot of **key enablers and cross-cutting factors** that empower progression along the pathways, such as engineering skills and tools and also the skills that are required for operating and managing the digital systems and tools. Another example is the many technological building blocks that need to be put into place and need to be integrated: from data communication infrastructure such as fieldbuses, industrial wireless or cabled networks to data storage, simulation tools, high performance computing, big data technologies, artificial intelligence, etc. The ConnectedFactories project endeavours to cover all, or at least the most important, of these different aspects, which are referenced as a self-standing **structured glossary** (available from the structured wiki[9] of the EFFRA Innovation Portal).

Every pathway involves the interaction of many stakeholders: one or more manufacturing companies, technology suppliers, system integrators, etc. The **pathways are also co-existing and non-exclusive**. In other words: the same manufacturing company can associate itself to different pathways and might be positioned on a more advanced level within one pathway compared to another. In addition, the same digital tools (for instance, digital platforms) may be applicable within different pathways and hence the suppliers of digital solutions, or system integrators, may be associated to more pathways as well.

Different people within companies will look at the pathways from their particular angle. A production manager would probably focus on manufacturing operations, a supply chain manager would look at value networks and a customer service manager would look at product service issues. Typically, the CEO (or CIO) of the company would, or should, look at all aspects related to their company strategy.

The most advanced level in the pathways is not necessarily the desired level for every manufacturing company. The pathways are not developed in order to distinguish 'good' manufacturing companies from 'bad' manufacturing companies. On the contrary, it is of particular importance that **within each and every pathway, the perspective of the small scale (SME) manufacturing**

---

[9] https://portal.effra.eu/wiki

companies should be highlighted. It should be noted however that large manufacturing is also facing challenges in making their way along the pathways.

The ConnectedFactories CSA has concluded its journey at the end of 2019, and the latest evolution of the three pathways in the perspective of Big Data Technologies adoption can be seen in the Figures below.

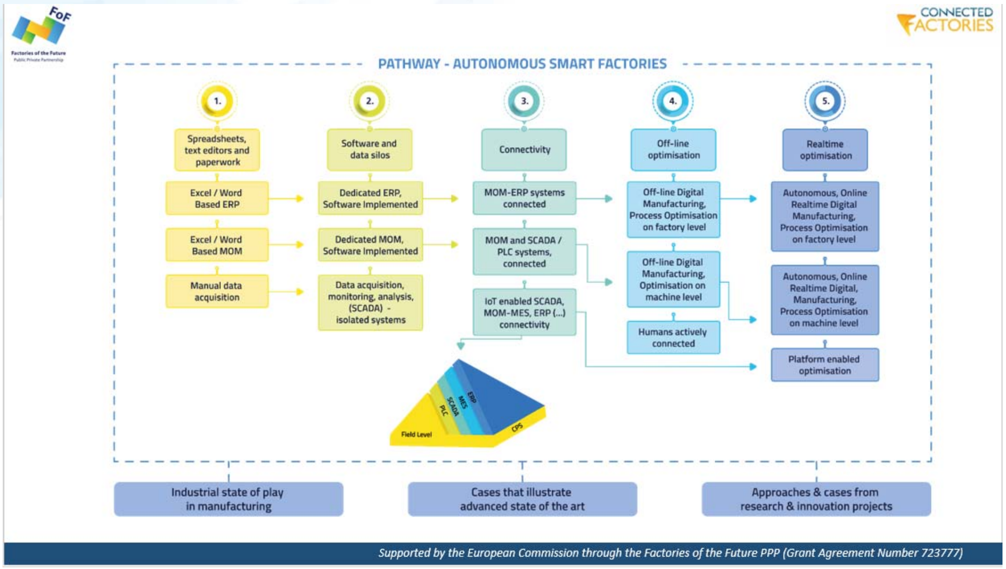


Figure 5 Autonomous Smart Factories

With respect to value creation resulting from big data, the Autonomous & Smart Factories pathway provides fundamental milestones that need to be reached in order to generate the data (in particular on level 3). Big amounts of data can then be used for the optimisation of manufacturing processes within the factory (still within the Smart & Autonomous Factories pathway). However, this optimisation will increasingly involve services that are provided by the suppliers of manufacturing systems or specialised third parties and described in the Collaborative Product-Service Factories pathway (depicted below).

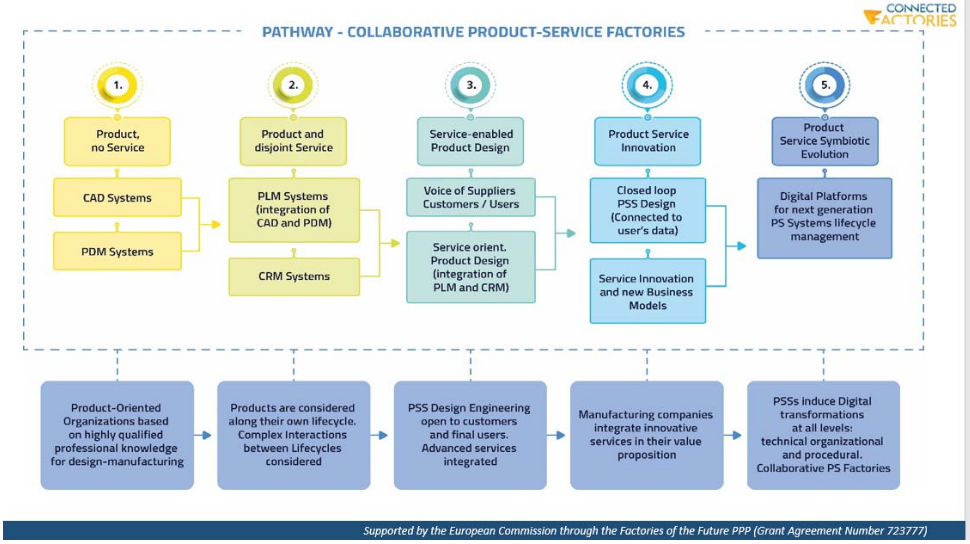


Figure 6 Collaborative Product-Service Factories

The hyperconnected Factories pathway is perhaps less directly associated with big data, unless large amounts of data are used to establish advanced brokerage process (hence supporting the engagement with existing and new value network partners).

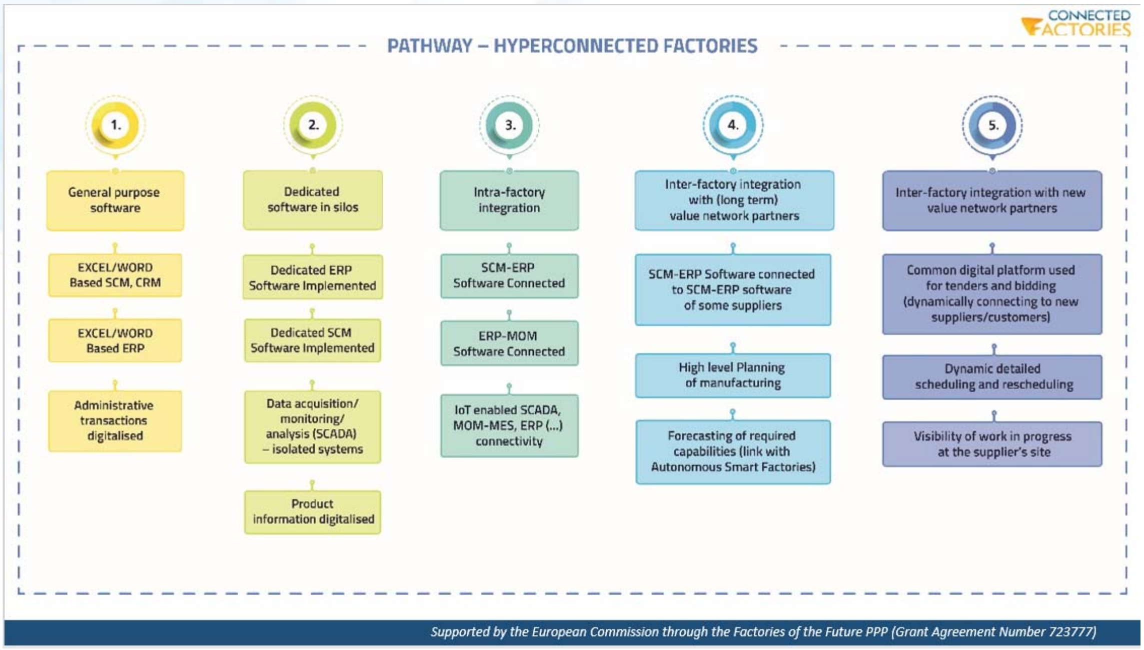


Figure 7 Hyperconnected Factories

In December 2019 a new ConnectedFactories 2 CSA was launched in the domain of Digital Manufacturing Platforms and is closely associated with the projects funded under DT-ICT-07 in Manufacturing.

3.1.2 A data-driven Architecture for next generation Digital Shopfloor: the DSA initiative

The AUTOWARE, Deadalus and Far-Edge projects under the Digital Shopfloor Alliance initiative set the ambition to develop a Service Reference Framework (RF) fully compliant with RAMI 4.0 and IRA reference metamodels and that will comply with ISO/IEC/IEEE 42010 Reference Architecture representation (functional, information, networking and system deployment view). The DSA data-driven autonomous service reference framework encompasses Industry 4.0 cognitive manufacturing technical enablers, i.e. data-drive and robotic AI systems, smart modular machines, cloudified control, secure cloud-based planning systems and applications to facilitate gradual and incremental development and deployment of cognitive automation systems while exploiting cloud technologies and smart services. This initial effort is now supported by 4ZDM and ForeSee cluster projects (zero defect and predictive maintenance) as well as the recently started LEVEL-UP project. Moreover, the data-driven architecture for next generation Digital Shopfloor Reference Service Framework for Software-Defined Automation & Control has a broad industrial capability, maps applicable technologies to different areas (IT/OT) of the enterprise, and serves as a guide for the deployment of Industry 4.0 technologies, smart digital manufacturing platforms and services supported by open international standards. The general Digital Shopfloor Reference Service Framework is maintained by the Digital Shopfloor Alliance (DSA).

The main objective of the DSA is to provide a reliable, cost effective integrated platform to provide solutions and services to support European enterprises of any size, in terms of both customized and flexible applications. Projects involved in the DSA under the coordinated support of the EFFRA

ConnectedFactories initiative and OpenDEI cross-sectorial platform coordination, provide a complete CPPS (Cyber-Physical Production System) solution allowing manufacturing and ICT industry to access all the different components in order to develop digital automation cognitive solutions for their manufacturing processes.

The DSA has proposed a service reference framework for autonomous services that consists of four main pillars (modelling, digital services & platforms, digital infrastructure, cybersecurity) and four layers/levels (as can be seen in the Figure below), which target all relevant layers where digital service platforms can be deployed to cover the whole manufacturing process from the shopfloor to the cloud. Moreover, those four layers of the architecture organize all the components/applications of an enterprise in their corresponding architecture layer depending on the intended functionality and smart service workflow. The layers are fully aligned with those proposed by RAMI 4.0 IEC 62264 specifications.

The digital manufacturing platform & service operation pillar can be divided into various layers as follows:

- **Enterprise.** The enterprise layer is the top layer of the DSA reference service framework and encompasses all IT enterprise services connected to the Plant IT network. That connection is supported by wired and 5G wireless connectivity. This includes Engineering, AR/VR and Business Management services.

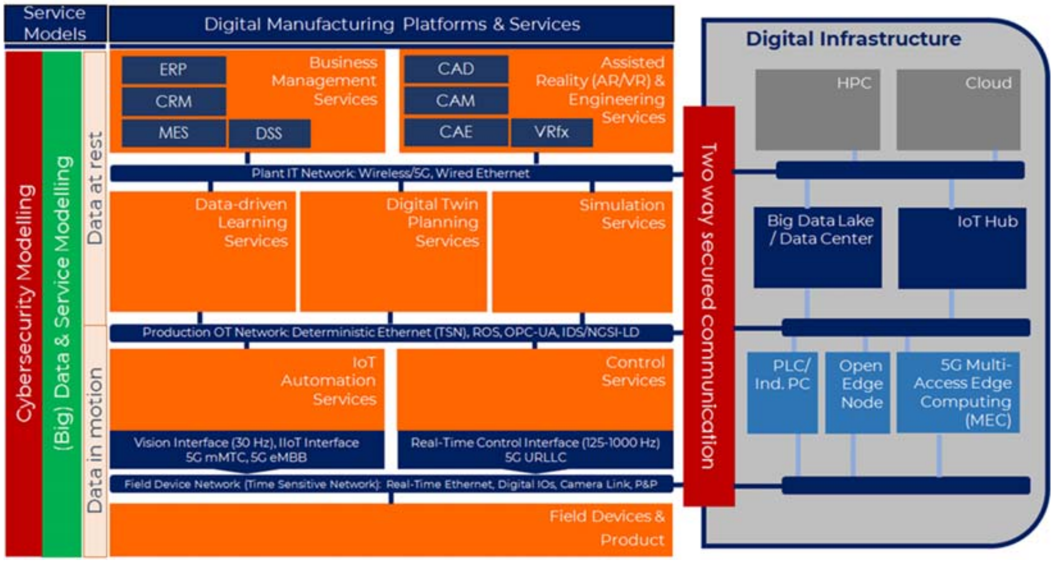


Figure 8 DSA Reference Service Framework focused on the Digital Infrastructure Pillar

- **Plant.** At the plant layer, a single factory is depicted and the layer is responsible for services dealing with insight generation and medium-long term data-driven planning and forecasting. This includes all the various workcells or production lines available for the complete production. This layer is connected to the services needed to holistically manage production. Therefore, this layer addresses data-driven learning services, digital twin based planning services and First Prime simulation services. The Factory Layer acts as a bridge between IT applications and OT network operations, hence, these services represent a middleware across OT and IT environments.
- **Workcell/Production Line.** The workcell layer represents the individual production line or cell within a company. Nowadays, a factory typically contains multiple production lines (or production cells), where individual machines, robots, etc. are located. Therefore, this layer refers to services mainly addressed to the management and operation of such working environments including IoT Automation and Control services, leveraging on edge-powered

control functions and advanced services for automated and increasingly autonomous sensing & actuation.

- **Field Devices.** The field devices layer is the lowest level of the reference architecture, where the actual machines, robots, conveyer belts, controllers, sensors and actuators, etc. are positioned. This layer is also the one where the actual product is placed. Embedded services related to the control and operation of the individual machines and manufactured products are placed in this layer.

The four layers allow the implementation of advanced analytics, simulation and automation services in a modular and software-defined manner. The four layers are connected by three main pillars:

- **Digital Service Modelling Pillar,** focuses on the modelling (service, data, AI models and vocabularies) of the different technical components inside the different layers (green column in Figure above). On each layer, different tools or technologies are applied and for all of them different modelling approaches are available.
- **Cybersecurity.** This pillar is focused on offering a reliable and secure use of the communication and information technologies. This safe and reliable environment should be offered through all layers of the company, from the plant to the cloud.
- **Digital Infrastructure Pillar.** This pillar is intended for the Fog/Cloud/HPC infrastructure required for the operation of the digital services pillar as well as wired (e.g. IEEE 802.1 TSN) and wireless communication (5G, WiFi, Zigbee…) and data distribution enablers, to create direct interaction between the different layers. This layer is, therefore, focused on the enablers for (big) data ingestion, processing and management for both data in motion and data at rest.

3.2 Validating and Prioritizing BDVA Technical Challenges in SMI

In the 2018 edition, the relationship between the BDVA Technical Challenges and the three ConnectedFactories 2025 pathways resulted in the identification of 56 challenges to be addressed by joint effort between the Big Data and the Manufacturing "Research and Innovation" communities, in order to keep the European Manufacturing industry competitive in the global marketplace.

Thanks to the active cooperation of several H2020 projects in both areas (Big Data and the Manufacturing "Research and Innovation" communities) and to the strong contribution of our SMI community (in BDVA), we have firstly grouped the 56 challenges in to three clusters (per grand scenario) and then proceeded with a polling session (physical and virtual) to understand the priorities and relevance of the challenges in view of the next Framework Program of the European Commission. The results of such a synthesis and prioritisation effort is reported in the sections below for Smart Factory, Smart Product and Smart Supply Chain scenarios (SF, SP, SSC) while lessons learned and recommendations are collected in Chapter 6.

Next to each challenge name, in brackets, is reported the reference code for those 56 challenges identified in the 2018 version of the Discussion paper. The related level of priority is also reported at the end of the challenge description.

3.2.1 Smart Factory Scenario

Smart Factory Context-driven Data Integration (SF8+SF1+SF3). Context-driven combination, integration and annotation (on-the-fly or stored in its natural format) of heterogeneous sources of data coming from legacy devices, CPS and IT systems (ERP, MES). This is needed in order to take advantage of the advanced processing capabilities (including compression and streaming) of new embedded hardware that generates data-in-motion (at machine or sensor level) together with data-at-rest coming from upper manufacturing levels (cells, production lines, sites) to perform such combination, integration and annotation. PRIORITY HIGH
Smart Factory Optimal Analytics Infrastructure (SF5+SF6+SF7). Some organizations, especially those dealing with sensitive data, are reluctant to put it on the cloud. They need to decide if data pre-processing tasks should be done on-premises or in the cloud. With traditional cloud and HPC

automation architectures, it is very difficult to know exactly where data is stored, so that until technology overcomes this hurdle, they prefer the on-premises solutions and edge nodes to process data close to the manufacturing point. In addition to that, industry 4.0 technologies like IoT and open interfaces (and APIs) imply a non-traditional way of data source vertical integration (from process level 0 to Planning and logistics level 4 – IEC 62264) that further complicates the management of analytics infrastructure. PRIORITY MEDIUM

Artificial intelligence and Factory Knowledge extraction (SF13+SF10+SF11+SF12). Whether it is Deep Learning or other Machine Learning techniques, it is expected that such technologies will enhance the ability of workers to interact with data and machines. This will open the door to support advanced decision-making at various factory levels, such as diagnosis of a machine failure, product quality control or business decision making through advanced prescriptive analytics of business KPIs (including estimate error/risk of predictions). It is necessary to make these technologies easy to use. PRIORITY MEDIUM

Secure Data Analytics at the Smart Factory (SF18+SF19). Data integrity is of key relevance as data is the basic information used, in particular, to make on-the-fly decisions. Yet data can be maliciously tampered with. Therefore, specific solutions to ensure data integrity should be included in analytics systems to provide robust data analytics solutions for industrial purposes. An example of such an integrity assurance solution is the Selective Anonymization in Smart Workspaces, but the optimal approach for which data has to be anonymized to support successful data analytics at site level, is yet to be investigated. Factory to Factory, such as B2B Data Exchanges, also imply trusted data networks to be built on top of Data Sovereignty principles. PRIORITY HIGH

Natural Language Interaction for Smart Factory decision makers (SF21+SF22+SF23). Detected patterns in data have to be conceptualized and different users (from novice to statistics experts) need to understand such patterns in the analytical process. It is needed to provide more seamless interfaces like natural-language interaction where workers ask for specific manufacturing information (manufacturing line performance) that needs to be extracted from data. These interfaces should be complemented with data exploration, navigation and annotation techniques to enable users to interact with the data effectively. Thus, the user will be able to reinforce machine learning at same time as getting continuous feedback from the system, thereby moving the stand-alone learning of the systems to user-machine bi-directional learning. PRIORITY LOW

As an immediate comment to this prioritisation, the majority of respondents consider that the **Data integration challenge** should have the highest priority. Today many companies struggle with extracting their data from the “siloes” created by different enterprise systems and functions (e.g. ERP, SCM, PLM, MES, CRM). Legacy Systems are usually strongly entangled with physical machinery and Cyber Physical Systems in a way that vendors’ lock-in severely affects and prevent cross-functional knowledge extraction and cross-domain development of advanced AI applications. The next highest priority was afforded to **Secure Data Analytics** along the computational continuum between the field, the edge and the cloud as a prerequisite to allow **AI and Knowledge extraction** (third challenge) or to define an appropriate **Analytics Infrastructure**. Finally, the lowest priority was given to the **Natural Language interaction** challenge in the Smart Factory. Indeed, human-centric interaction plays a fundamental role in next generation workplaces in the smart factory, but it is possible that the presentation of the challenge was somewhat misleading, perhaps envisaging scenarios of colloquial NL interaction between humans and machines which were too futuristic for a noisy and demanding factory environment.

3.2.2 Smart Product Lifecycle scenario

Product, Data and Service Interoperability (SP1-2-3-10). EU competitiveness in the global market requires not just the manufacturing of high quality products, but also the provision of value added personalized services to customers. Data generated by Smart Intelligent products (basically IoT real time data) needs to be interoperated with Smart Intelligent services (basically AI-driven) in a symbiotic way. Product marketplaces are not enough any longer and advanced multi-sided data-service marketplaces need to be integrated. This is particularly true for complex products such as robots, industrial equipment, aircrafts, ships and vehicles. PRIORITY HIGH

Product Lifecycle Data Models and Asset Administration Shells (SP4-5-6-10). Product Data Models have existed for decades, especially to support collaborative distributed product design and engineering (PROSTEP IVIP standards). Virtual Obeya[10] environments support today the co-operative design work of engineers through continuous and secure Product design Data Sharing. The manufacturing industry of the future aims at modelling its products not just in the design / engineering phase, but through the whole lifecycle and calls for different data models to be integrated: e.g. as designed, as manufactured, as maintained, as recycled models are usually created and stored in different administration domains. RAMI 4.0 and its Assets Administration Shell paradigm aims at creating a semantic interoperability infrastructure throughout the whole product lifecycle, as a data infrastructure for advanced services such as diagnosis, maintenance, optimization and in support of new business models as those dictated by Sharing and Circular Economy. PRIORITY MEDIUM

Secure Product Data Processing Architectures (SP7-8-11-12). Advanced AI tools and components are now embedded in smart products (e.g. not just robots and machineries but also cars, aircrafts and even shoes or garments), deployed in the proximity of the location of the product (Fog edge) and governed in private and public cloud data infrastructures. New data driven vertical interoperability architectures need to be developed in order to manage the complexity of such multiple and heterogeneous data sources. The IDS Reference Architecture Model is one way to represent the complex data sovereignty rules, which govern the data exchanges at the machine, edge and cloud level, encompassing real time Data in Motion and high- volume Data at Rest infrastructures. PRIORITY MEDIUM

Personal and Product Digital Twins integration (SP9-13-14-15-16-17). Digital Twin has now become a commonly used term, to indicate not just the static digital models of a physical object, but also its dynamic behaviour during operations. However, product operations also imply the more or less active presence of the humans who interact with it. In that case, user profiles are able to model the preferences, competences, roles and capacities of humans, but only in a fairly static way. The dynamic representation of humans is now called a Digital Persona - a sort of Personal Digital Twin. Modelling, simulating and optimizing the interaction between Products and Humans has now become a complex discipline, often called collaborative intelligence, implying “humans assisting products” and “products assisting humans” interaction. In the digital world, this means there is a need to integrate personal and non-personal digital twins in a common dynamic data-driven environment. PRIORITY MEDIUM

In this Smart Product (SP) scenario, business needs and requirements are driving the prioritisation process. SP is in fact perceived as the most disruptive and potentially revolutionary scenario for Manufacturing enterprises, allowing them to go beyond the factory, following their products along their whole lifetime and this way meeting the even stricter requirements of sustainability and circular economy of materials and components. This is the reason why all the four challenges received quite high priority scores. The most immediate business perspective is in **Product, Data and Service Interoperability**, where the service economy has already demonstrated its disruptive potential in many traditional industries such as automotive, aeronautics or machine tools. In order to ride the *servitisation* wave, an overall interoperability between Product, Data and Services is needed as well as the development of new standards which extends the Product Lifecycle to a Service Lifecycle (and a Data Lifecycle). Then, of course, new privacy preserving and data confidentiality/sovereignty **Secure Product Data Processing Architectures** are needed, alongside lifecycle Product Models which are able to integrate the different stakeholders of complex long living products (such as ships, aircrafts, machinery, but also cars and smart appliances) in interoperable **Product Lifecycle Data Models** and for the industrial assets in standard **Asset Administration Shells**.Furthermore, the human aspect plays an important role both on the professional side (engineers, workers, maintenance workforce) and on the citizens side, in terms of users, drivers, consumers (and sometimes prosumers). The **Personal and Product Digital**

---

[10] EU funded project - http://www.virtualobeya.com

Twins integration challenge foresees the lifetime integration between Things and Humans and opens the door to a new way of collaboration inspired by Artificial Intelligence, such as the Digital Assistants found in many modern cars.

3.2.3 Smart Supply Chain scenario

Smart supply chain vertical data integration (SSC1, SSC6). The integration, synchronization and standardization of data from different sources (i.e. order, customer, supplier, transportation, additional data such as traffic information or truck GPS data, plus data inside a company such as design, operational, product, financial data, etc.) enables data access to all companies across the supply chain. End-to-end vertically integrated and transparent supply chains will be possible by linking different multimodal and heterogeneous data, models and flows. PRIORITY HIGH
AI-driven supply chain models and services (SSC8, SSC10, SSC11). Avoiding problems such as overstocking, stock outages, delivery delays, etc. requires a balance between the autonomous, distributed processing of each entity in a supply chain and the collaborative, goal-driven coordination among them. The use of AI will improve the forecasting processes (e.g. demand forecasting, inventory planning, transportation management, etc.) and will open the door to self-optimization of a supply chain, aiming at a global optimum by considering all the connected organizations. PRIORITY MEDIUM
Human-driven Cognitive supply chain (SSC9, SSC11). Moving a product from supplier to customer requires a cross organization chain of extensive coordination between human and software systems. The ability to capture and understand the decisions made by humans (e.g. by monitoring their behaviour, what they said or wrote, etc.) and to make these decisions reusable in a similar context has the tremendous potential to disrupt the supply chain business. The cognitive services such as perceiving changes in a supply chain and knowing how to respond to these dynamic fluctuations will advance supply chains to the next level. PRIORITY LOW
Agile supply chain networks (SSC4, SSC5, SSC10). Manufacturing companies can leverage the opportunities arising from the globalized marketplace of capabilities and capacities. This will allow dynamic creation, reconfiguration and management of a supply chain network based on a company's profiles, loyalty, customer feedback, etc. as well as flexible sharing of production resources within the supply chain network. Typical marketplace services, such as searching for a partner candidate for a supply chain, can be enhanced with AI-based discovery of one or more partner candidates based on their capabilities, future availability, past performance, etc. PRIORITY MEDIUM
Traceable and trusted supply chain (SSC3, SSC13). To address increasing demands for provenance information and compliance across multiple industries, new technologies such as blockchain and industrial data spaces should be considered. Blockchain technologies will enable holistic and secure product tracing by giving confidence to customers concerning the origin, specification, producers, environmental safety, authenticity, etc. of a product. The IDS technologies (e.g. usage policy definition and enforcement) will ensure the protection of data sovereignty and trust in supply chains. Whereas strengthening traceability will mitigate the high costs of quality problems (e.g. reputational damage) and increased trust will create a better environment for sustaining supply chain relationships and for reducing risks and uncertainty. PRIORITY MEDIUM

The majority of respondents considers that the **Smart supply chain vertical data integration** has the highest priority, as interoperability has a high impact factor on the efficiency and on the costs within supply chains. Implementing data integration is the first step, before more intelligence can be added. It should act as an enabler for new service oriented business models. The respondents had different opinions on trust and traceability, mainly because the other pillars have first to be solved, in order to unlock the potential of the **traceable and trusted supply chain**. With recent advances in technologies, such as blockchain and IDS, trustworthiness and secured data exchange have become a very real possibility. The lowest priority was given to **the human-driven cognitive supply chain**, as the capturing rationale is hard to achieve. However, understanding implications and trade-offs could be a key competitive advantage in the coming years.

4 COMPLEMENTARY RESEARCH & INNOVATION AGENDAS FOR SMI

This section of the paper aims at addressing other complementary viewpoints within the "Big Data for Smart Manufacturing Industry" debate. In particular, we analyse in §4.1 the Industrial IoT perspective, thanks to contributions from the AIOTI and some of its working groups. In §4.2 the Robotics perspective and its "Agile Production" vision is explored. In §4.3 the Artificial Intelligence perspective[11] is considered from the demand viewpoint at field, edge and cloud layers.

4.1 Industrial IOT Standardization for SMI

There are several ways to visualise the existing IoT standardization landscape. The figure below shows a high-level overview of the global standardization ecosystem, where the left side, shows standards bodies that are formally recognized under the WTO, the EU Regulation 1025/2012, and national government contracts or rules, as well as so called fora and consortia.

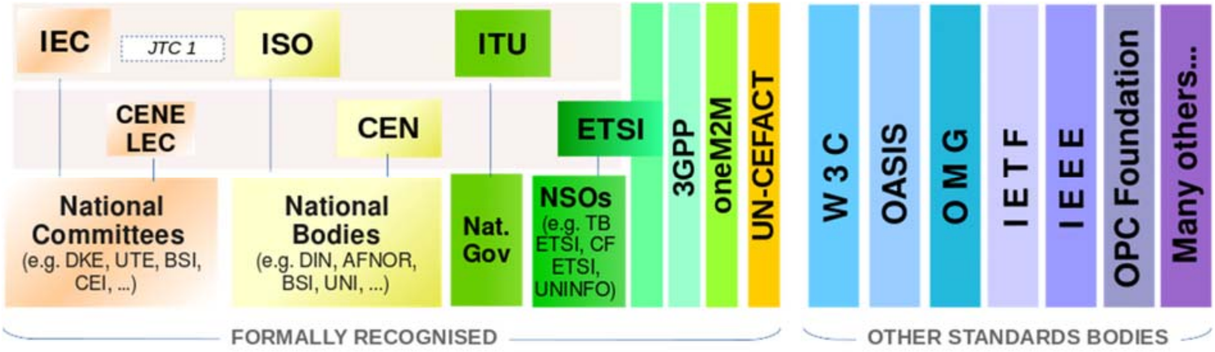


Figure 9 High-level overview of the global standardization ecosystem[12]

The formally recognized organizations are the IEC (International standardization organizations International Electro-technical Commission), ISO (International Organization for Standardization) and ITU (International Telecommunication Union), the ESOs (European standardization organizations) CENELEC, CEN and ETSI and as well, the standardization bodies on the national level.

Furthermore, although, the global partnership projects 3GPP and oneM2M are not formally recognized, both of them have ETSI as a driving member and thus some linkage into the formally recognized standardization environment. UN/CEFACT is a UN organization and thus also has some level of government recognition.

The right side of Figure 9 shows the organizations that are not formally recognized by governments. In particular, they provide technical specifications, usually in a dedicated area of expertise, for global use, and play a key role in the ICT sector where a number of relevant and widely implemented ICT standards are developed and maintained.

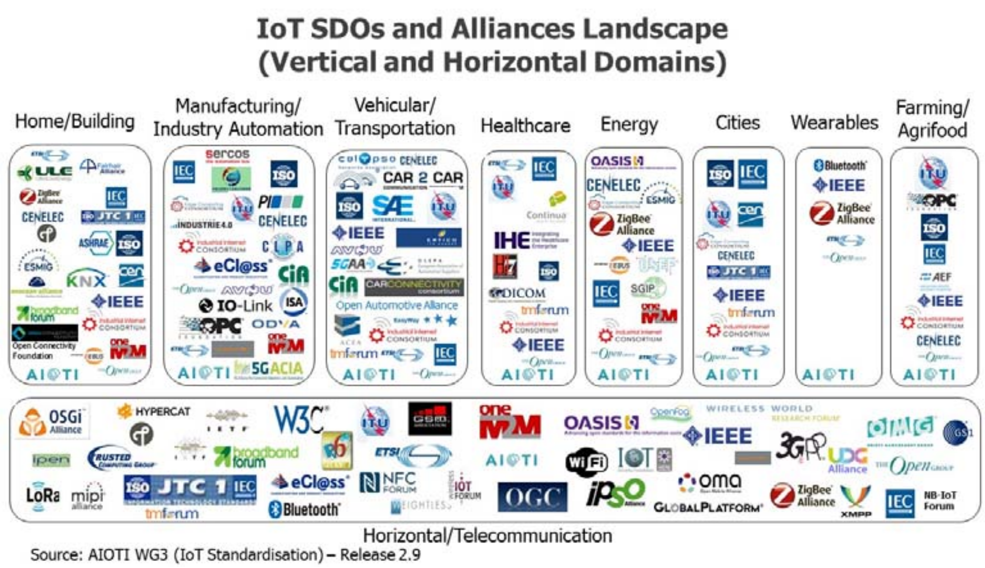
Figure 9 shows one of the IoT SDOs and Alliances landscapes derived by AIOTI on the global dynamics and landscapes of IoT related standards bodies, Alliances and OSS initiatives. These IoT

---

[11] A full "AI for Manufacturing" perspective is foreseen for 2020 edition of the paper, where the joint BDVA- euRobotics AI PPP SRIDA will be commented and interpreted on the basis of the demand side reported here in this document §3.2

[12] MSP/DEI WG on standardization - Task 2 findings

standardization landscapes can be used, among other purposes, to leverage existing IoT standardization, industry promotion and implementation of standards and protocols.



Source: AIOTI WG3 (IoT Standardisation) – Release 2.9

Figure 10 IoT SDO and Alliance Initiatives Projection on Vertical and Horizontal Domains[13]

The Multistakeholder Platform on ICT Standardisation (MSP) and Digitising European Industry initiative (DEI) joint working group (MSP/DEI WG[14]) identified **available standards and technical specifications in some critical technical areas:**

- **A critical technical area concerns the platform/service layer**, which constitutes the middle layer and is key for integration, as it creates the linkage between different organizational levels, in particular connecting the shop floor to the office floor, and interconnecting processes. In particular, two key platforms/service layer approaches available in the context of digitizing industry have been identified: **OPC Unified Architecture (OPC/UA) and oneM2M**.
- Another critical technical area that has been identified is a **Blockchain and other Distributed Ledger Technologies (DLT)**. Currently, there are several Blockchain-related standardization projects, including projects in ISO and ISO/IEC JTC1, ITU-T, CEN, CENELEC, ETSI, IEEE, as well as the Accord project.
- Other identified critical technical areas concern basic horizontal building blocks, security and privacy and matters specific to digitizing industry, the latter ones including (but not limited to) robotics, 3D printing, cognitive methods and Artificial Intelligence, web technologies, semantics, ontologies.
- It is important to emphasize that further work is required to consolidate the identification of digital industry specific standards.
- Particular attention has also been given to the identification of relevant **open source developments.** In many fields, open source developments are often accompanied by standardization activities flanking the open source developments or through the adoption of technologies from open source into standardization. In the other direction, open source is also a way for promulgating standards and promoting the implementation of standards.
    - EclipseIoT, OpenAAS and OPC-UA Open Source have been identified as open source platforms and projects relevant to the digital industry (and IoT).

---

[13] *AIOTI report "IoT LSP Standard Framework Concepts" (new version (2.9) not yet published)*

[14] https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=57546

For details on the **available standards and technical specifications mentioned in this section, please see** Task 2 findings[15] of the Joint MSP/DEI WG on standardization.

Recommendations towards EU Rolling Plan on ICT Standardization for 2019

The Joint MSP/DEI WG recommended nine high-level actions to be included in the EU Rolling Plan on ICT Standardization for 2019. A summary of these nine actions is provided below:

- **Action 1:** Common communications standards and a reference architecture for connections between machines (M2M) and with sensors and actuators in a supply chain environment are a basic need and a priority.
- **Action 2:** As part of the new skills agenda for Europe, European Standardisation Organisations (ESO) could check whether the e-skills standards sufficiently account for the manufacturing skills of Key Enabling Technologies (KET), including future manufacturers, M2M, rapid prototyping and others.
- **Action 3:** Conduct a study to identify and analyze opportunities for revisions of existing standards (communications, M2M) or new standards with a particular view on new production technologies, manufacturing processes including lifecycle operations (circular economy), functional safety issues and skills-deficit reduction.
- **Action 4:** Improve interoperability and reduce overlap, redundancy and fragmentation. Often there are several standardization activities ongoing in the same area in parallel.
- **Action 5:** Interoperable and integrated security - SDOs should work on interoperability standards for security and for linking communication protocols in order to provide end-to-end security for complex manufacturing systems including the span of virtual actors (from devices and sensors to enterprise systems). Standards should take into account risk management approaches as well as European regulation and regulatory requirements.
- **Action 6:** Create a hierarchical catalogue of technical and social measures for assuring privacy protection and task all SDOs impacting the DEI domain in general and the advanced manufacturing domain in particular to comment on and prioritize the elements in the catalogue
- **Action 7:** Standards should be developed to define the main characteristics for all levels of the interaction from mechanical to electrical to protocol to semantic levels between robot and tool to ensure the exchangeability and to enable the design of generic tooling (plug-and-play). There are 2 main types of End Effector: "Off-the-Shelf" and "Bespoke". It is desirable that off-the-shelf end effectors operate on a single software protocol. There is a need for Industry 4.0 to standardize this. It would then become Plug-&-Play. For "Bespoke" end effectors (most commonly purchased) the system integrator specifies the software protocol for the Robot and End Effector.
- **Action 8:** Start the discussion about the possible development of harmonized standards in the area of additive manufacturing. Currently, there are no harmonized standards under the Machinery Directive for Additive Manufacturing (AM) equipment.
- **Action 9:** Develop standards for ensuring long-term traceability of materials, to enable re-use and recycling.

4.1.1 Industrial Internet Consortium IIoT Connectivity Model

The **Industrial Internet Consortium** (IIC) is the world's leading organization with the mission to deliver a trustworthy IIoT, where systems and devices are securely connected and controlled to deliver transformational outcomes. To achieve this, the IIC delivers best practices, reference architectures,

---

[15] http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=53617

case studies, and standards requirements to ease deployment of connected technologies, therefore influencing standards development processes relating Internet and industrial systems.

A key point that the IIC aims to enhance is connectivity. Connectivity is one of the foundational technologies as it provides the ability to exchange data amongst participating components of an IIoT system. In industry, connectivity facilitates an improvement in efficiency, productivity, quality and safety of processes thanks to real-time monitoring and management. However, the classic automation architecture and the heterogeneity that is found in the plant, full of proprietary technologies and specialized connectivity standards optimized for domain-specific use cases, can present a challenge in achieving these objectives.

To address this issue, the IIC introduced its IIoT Connectivity Stack, which defines an open reference architecture to enable data sharing and interoperability between previously closed components and new applications.

Figure 11 shows the proposed IIoT connectivity stack, which is derived from the 7-layer model and the Internet 4-layer model, to fit industrial and IoT connectivity requirements, by adding the new framework and distributed data layers.

On the right side of Figure 11, the 'core' protocols for each layer are shown. These core solutions are the selected solutions considered the most relevant and appropriate to fulfil connectivity and communication requirements in industry use cases, established by IIC criteria.

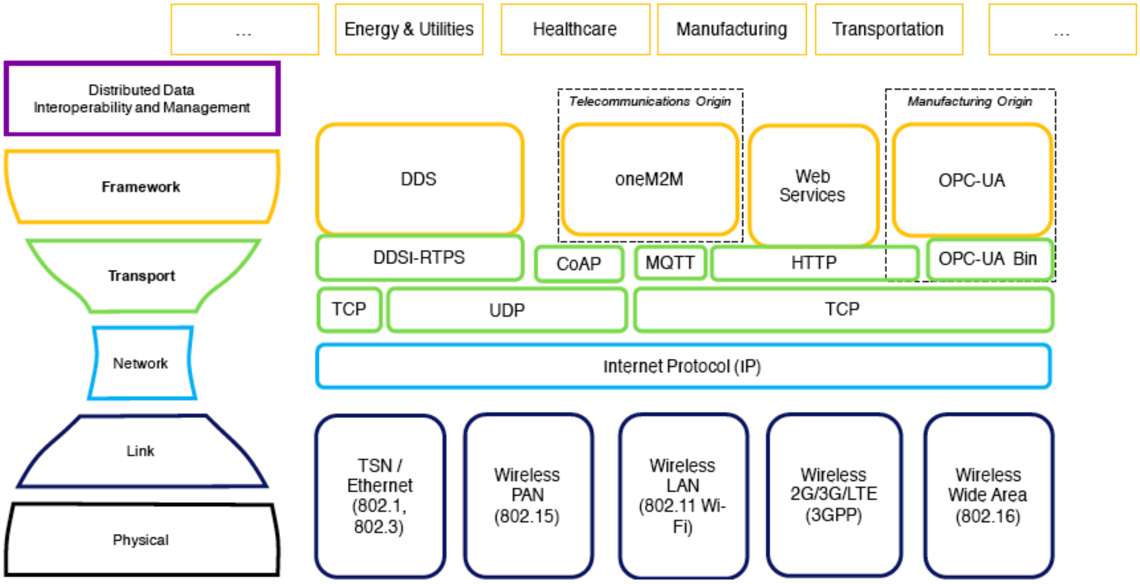


Figure 11 IIoT Connectivity Stack Model. Source: Industrial Internet Consortium Connectivity Task Group

While connectivity refers to the infrastructure enabling communication between elements, communication refers to the information exchange between them. Communication is the basis for interoperable systems, and all these different components and systems need a minimum common context to be able to communicate meaningfully. To achieve interoperable communications among endpoints, three main aspects should be considered: technical interoperability, syntactic interoperability and semantic interoperability.

Regarding the new layers defined in the IIoT stack, the top Distributed Data and Interoperability layer is in charge of providing the semantic interoperability, this is, the ability to recognize unambiguously the information in the exchanged data for different contexts.

The connectivity framework layer is responsible for providing and handling unambiguous structured data exchange and therefore syntactic interoperability between participants, regardless of the platform or the programming language. This layer is therefore essential to build IIoT systems and applications.

Technical interoperability requires defining unambiguously the exchange of information as bits and bytes between two endpoints, assuming that the information exchange infrastructure, formed by the physical, link and network layers are established, and allow this exchange. Therefore, the main responsibility for the technical interoperability falls upon the transport layer.

The Physical and Link layers are much more dependent on the requirements for the communication channel, going from wired to short and long-range wireless solutions, supporting a wide spectrum of technologies, which are valid as long they support an IP Network layer.

IP allows for the creation of address spaces for the networked elements and routing data from source to destination hosts even on different physical links, which is the cornerstone and heart of any Internet and IoT application. This relevance of IP as the central pillar of the connectivity model is represented in the shape of the IIC proposed layered stack, which is not arbitrarily hourglass shaped. While for the rest of the layer, the protocols and technologies to choose from are many, for the network layer the only considered option is IP, being the narrower part of the stack.

Regarding the transport layer functions, they include endpoint addressing (as several endpoint may coexist in the same IP address), modes of communication (unicast, multicast, and broadcast), network topology, connectedness (TCP or UDP), policies, synchronization, and message security.

The choice of protocols for this transport layer also comprises the selection of a communication model or paradigm that falls into two categories: publish-subscribe (MQTT, AMQP, DDS are popular implementations) or request-response (HTTP, CoAP), each of them with pros and cons, which are mainly due to complexity, scalability, or a/synchronicity requirements.

As can be seen, IIC allows for multiple core connectivity standards, and considers that deployments will vary per system unique requirements. The proposed model provides a reference so IIoT architects can perform their own connectivity technology evaluation and selection in order to provide solutions that are interoperable and fit within business and industrial operations.

4.1.2 Data-driven Smart Manufacturing standardisation challenges

This subsection provides a list with key recommendations that are specific to data-related Smart Manufacturing standardisation.

Data access and data sharing:

- Global standards e.g., ISO/IEC JTC 1 SC 42 are needed to establish a framework for exchanges;
- Need to integrate semantics and metadata into the ecosystems, where metadata interoperability can be considered as a gap;
- Data models: developed on a proprietary basis and mostly specific to the vertical domains to which they apply; Property-based systems, as used in e.g., oneM2M and IEC Common Data Dictionaries (e.g. IEC 61987) need to be developed further;
- Ontologies: need to enable a greater interoperability between systems (and a certain de-layering of the manufacturing pyramid) will foster the evolution of the existing (common) data models towards a more systematically developed and managed set of ontologies;
- Data Governance: need to define rules on data storing, data access and data sharing.

Data protection/privacy:

- The fitness for purpose of existing standards that relate to Smart Manufacturing scenarios must be clarified. In the case of automated communication across domain boundaries (such as the boundaries between jurisdictions), the relevant data protection requirements and associated security requirements derived from these must be harmonized.

Technical negotiation of (capability-based and property-based) **security profiles** for Smart Manufacturing:

- Need standards for the technical negotiation of (capability-based and property-based) security profiles for Smart Manufacturing communication and/or cooperation of entities within a variety of security domains. This includes:
  - Technical support for the classification of information, and requirements for the handling of suitably classified data;
  - Identification and authentication requirements;
  - A method of evaluating the trustworthiness of cooperation partners.

Trustworthiness of the value-added network:

- Define process standards for the security of the cooperation within the value-added network. This includes a method of evaluating the trustworthiness of cooperation partners. Typical mechanisms include manufacturers declarations, certificates, auditing:
  - Rules for the sharing of classified data and information;
  - Minimum requirements regarding security for B2B.

Standardized concept of roles and permissions for parties involved in Smart manufacturing:

- Access to data and resources in the context of Smart Manufacturing cooperation necessitates standardized rules. Existing concepts, such as IEC 62351, can serve as a starting point. Boundary conditions governing implementation include scalability and the potential for representation in the form of specific vertical requirements.

Security infrastructure for safe inter-domain communication:

- Secure communication requires secure identities (identifying factors and attributes) and trust anchors. Generating and administering secure identities and securing their trustworthiness requires a secure infrastructure. The requirements for this include factors such as scalability, resilience, cost-effectiveness, long-term fitness for purpose, and (user-defined) trustworthiness beyond, and independent of, local jurisdictions.

4.2 Robotics Research and Innovation Challenges for Agile Production

Due to rising customers' individual demands and rapidly changing markets, agility has become a primary condition for competitiveness in the Manufacturing industry. To cope with such variety and volatility of demands, a transition towards higher levels of agility in the production processes is needed.

In the future of agile manufacturing, humans and machines work closely and efficiently together within the production and operational processes, for example on the shop floor. Specifically, **human-robot collaboration (HRC)** is expected to become an important enabler of agile production. The vision is of a robotic system which is simple to program (e.g. by workers on the shop floor that give examples to the robot), which is able to accomplish variable tasks with variable and shared autonomy in collaboration with the worker and enhancing their cognitive abilities rather than replacing them. Both Artificial Intelligence (AI) and Big Data (as the basic resource) are two key ingredients to build higher levels of autonomy and adaptation in such collaborative robotics.

The future AI PPP will connect and coordinate European AI, Big Data and Robotics innovation and research communities in support of the implementation of the European AI strategy[16] and specifically, of the first of the three pillars of the European approach to AI[17]:

- **Boosting the EU's technological and industrial capacity and AI uptake across the economy**, both by the private and public sectors (national and EU levels), including a joint coordinated effort for investment in research and innovation, and access to data.
- **Prepare for socio-economic changes** brought about by AI by encouraging the required changes in education, training systems, labour market, and of social protection systems.
- **Ensure an appropriate ethical and legal framework**, based on the Union's values and in line with the Charter of Fundamental Rights of the EU.

Regarding the first pillar, the EU seeks to stay at the forefront of technological developments in AI and ensure they are swiftly taken up across its economy to enhance the competitiveness of European companies. This implies, among other things, increasing investment to strengthen fundamental research and make scientific breakthroughs, upgrading AI research infrastructure and facilitating access to data. Moreover, this requires a joint effort by both the public and private sectors to gradually increase overall investment by 2020 and beyond. Without such efforts, the EU risks losing out on the opportunities offered by AI, facing a brain-drain and being a consumer of solutions developed elsewhere.

In 2019, the **euRobotics** (European Robotics Association) and the BDVA, released a common vision paper for an AI Public Private Partnership (AI PPP), including a research, innovation and deployment agenda. This was specifically “directed at boosting the industrial uptake of AI and ensuring Europe’s world-wide leadership in developing and deploying value-driven trustworthy AI based on European fundamental values.”[18] This partnership is envisioned to embody a close relationship between the (cyber-)physical (robotics) and data worlds, and it is expected to bring better decision-making, more decisional autonomy, and improved human-machine interfaces, all of which are key elements for successful HRC.

In that common vision paper (*Strategic Research, Innovation and Deployment Agenda for an AI PPP*), euRobotics and the BDVA identified a series of technology enablers, in the “***Cross-Sectorial AI Technology Enablers***”[19] section, which are critical for any AI-based robotics application, specifically for intelligent automation with variable autonomy through robotics and human-robot teaming in the manufacturing industry. In the following paragraphs we summarise the specific challenges identified for each of those enablers:

Sensing, Measuring and Perception. Sensing, measuring and perception technologies are fundamental to building digital representations of the physical environment, where physical motion and the inner self are both fundamental for robots to acquire situational awareness of the world around them. This is critical for successful robot decision making, control, and learning in dynamic, uncertain, complex environments. Measuring means to enhance sensing thanks to calibration (to frames of reference), while perception means to model the extracted information into data assets that can be utilised by the robot. Some specific challenges are: the development of faster more accurate and complete methods of perception (capturing multiple data modalities, including their fusion), which are reliable across changing operational conditions, such as weather and light, or human genders, ages and ethnicities; methods for *active perception and learning*; modularisation and standardisation of sensor interfaces; perception methods which can adapt to changing operating conditions or dynamics, including available resources (e.g. power); sensor systems which are resilient to challenging environments; methods to validate and certify them, as well as the perception processes based on them, for safety, privacy, trustworthiness; and self-calibrating sensor systems.

---

[16] https://ec.europa.eu/digital-single-market/news-redirect/624220

[17] European Commission, 2018. “Artificial Intelligence for Europe,” COM(2018) 237 final, Brussels, 25.4.2018.

[18] http://www.bdva.eu/AIPPP-Vision-paper-PressRelease

[19] Strategic Research, Innovation and Deployment Agenda for an AI PPP: A focal point for collaboration on Artificial Intelligence, Data and Robotics. Second Consultation Release, September 2019. A joint initiative by BDVA and euRobotics.

Continuous and Integrated Knowledge. This involves continuously transforming, cleaning, storing, sharing, modelling, simulating and synthesising data, and extracting insights through a combination of data-driven and knowledge-based models. This allows for the full automation of the perception-decision-action loop with means to guarantee the reliability and safety of the in AI functionality. The main benefit this combination provides is *transparency* (and the subsequent *explainability*) of AI systems, critical in safety-critical scenarios such as self-driving cars and drones. It also allows for keeping truthful digital twins representations of products, processes, the robot itself, or any other physical entity on the shop floor. Some specific challenges are: the *scaling and federation of robotic systems; methods for data augmentation* and *data synthesis* when it is too difficult, costly or unsafe for the robot to collect the data (e.g. experience, as in reinforcement learning); methods for *knowledge modelling* and representation that enable the seamless integration of data and connection with the physical world; *methods* to ensure scalability and reusability of analytical outcomes, such *as transfer learning*, *online (e.g. continual lifelong) learning*, *meta-learning* and *knowledge representation learning*; methods that *integrate data-driven and knowledge-based approaches* to guide the learning process or *to ensure transparency of models learnt from data* to ensure compliance to given specifications; and the development of methods for handling *security and privacy concerns* regarding robots collecting data in human environments.

Trustworthy Hybrid Decision Marking. On the factory floor, we can envision future human-robot collaboration in which the responsibility for the perception-decision-action loop is shared between worker and robot. Several scenarios have been identified: *Human Decision Making*, *Machine Decision Making*, *Mixed Decision Making and Decision Support*, and *Sliding or Variable Decision Making*. In the last case, the balance between human and robot decision making varies during operation depending on robot confidence levels or human interactions. At this level, the *human agency* problem arises, which required keeping the workers in control, ensuring the collaboration is beneficial to them, enhancing their capabilities rather than replacing them. These scenarios are more or less affected by core challenges such as timeliness, robustness, and trustworthiness, resulting in the following specific challenges*: interpretation of context, dealing with uncertainty*, *transparent anticipation, reliability*, *human-centric planning and decision making.*

Physical and Human Action and Interaction. To achieve a seamless operation in unstructured, dynamic, uncertain and ambiguous environments with people, robots need to work in harmony with their social, physical and environmental context. There is a set of core challenges related to the processing of environmental cues to guide the decisional autonomy that drives the sequences of individual actions that form an interaction. This can involve multiple sources of data and their fusion prior to their analysis (e.g. interpretation) within the context of an interaction sequence. For example, interpreting the meaning of the spoken word in the context of an on-going interaction, or unexpectedly detecting liquid in glass and the effect on an on-going grasping plan. Some specific challenges are: seamless and *natural interaction in unstructured contexts* (including multi-modal interaction), *improved natural language understanding and dialogue;* development of verbal and non-verbal interaction models (including gesture and emotion-based interaction); *development of interaction technologies using Virtual Reality (VR) and Augmented Reality (AR), to ensure safe interaction in safety-critical scenarios*, *confidence measures* for interaction.

4.3 Artificial Intelligence for Manufacturing Industry

The rapid progress of AI technologies is also impacting Smart Industry areas. On 19th February 2020, the Commission unveiled its vision for a digital transformation that works for everyone. This includes a White Paper[20] proposing a framework for Artificial Intelligence based on excellence and trust. One particular challenge identified by the Commission is the extremely fast development of AI. Simply put, AI is a collection of technologies that combine data, models, algorithms and computing power. Advances in computing and the increasing availability of data are therefore key drivers of the current upsurge of AI. Europe can combine its **technological and industrial strengths with a high-quality digital infrastructure and a regulatory framework** based on its fundamental values to become a global leader in innovation in the data economy and its applications as set out in the European data strategy[21].

---

[20] https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

[21] https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf

Smart manufacturing is not disconnected from this vision and actually the White Paper refers to a new generation of products and services in areas where Europe is particularly strong (machinery, transport, cybersecurity, farming, the green and circular economy, healthcare and high-value added sectors like fashion and tourism). Moreover, Europe should drive their efforts towards **ecosystems of excellence** (intended to develop advanced -digital- value networks) and **ecosystems of trust** (human centric AI and digital sovereignty).

In the next years we are set to observe an **even greater explosion of data being generated in the manufacturing domain and a remarkable change in the proportion of computing tasks that are performed across smart connected assets, edge and cloud.** In fact, the volume of data produced in the world is growing rapidly, from 33 zettabytes in 2018 to an expected 175 zettabytes in 2025. Moreover, the current computing distribution that today allocates 80% of data processing and analysis to the cloud (data centres and centralised computing facilities), and 20% to smart connected objects, such as cars, home appliances or manufacturing robots, and in computing facilities close to the user ("edge computing"), is also set to change dramatically.

With this framework in mind, and the challenge of contributing to the development of ecosystems of excellence and trust in Smart Manufacturing Industry, with an expected explosion in data generation and a very different data storage and processing component, AI in manufacturing will also observe an increased distribution and cooperation in AI services across layers of increasingly smarter and connected assets, workers, customers, operations, products and manufacturing and supply chains. Hence, this SMI group is proposing the framework illustrated in the Figures below to address and analyse the challenges for AI technologies in machines (§4.3.1), edges (§4.3.2), and Cloud (§4.3.3).

The AI challenges can be observed in layers where AI should make decisions on increasingly larger time scales. AI in manufacturing will generally face challenges in terms of the velocity (decision speeds from milliseconds to days-months) and the volume, veracity and variety of data that would need to be processed. Therefore, AI should be distributed across the automated shopfloor, edge, data spaces and service layers (in the cloud) to support the operation of business processes in the autonomous factory. This implies that we should leverage the technology enablers for distributed edge AI operations with cognitive gateways that exploit trusted and sovereign data ecosystems for AI service development and operation; i.e. industrial dataverses. In a highly distributed and dynamic environment the development of highly modular pluggable and efficient AI services (algorithms, tools and models) is key to ensuring the digital continuity of manufacturing operations and the seamless alignment of digital twin and physical twin operations.

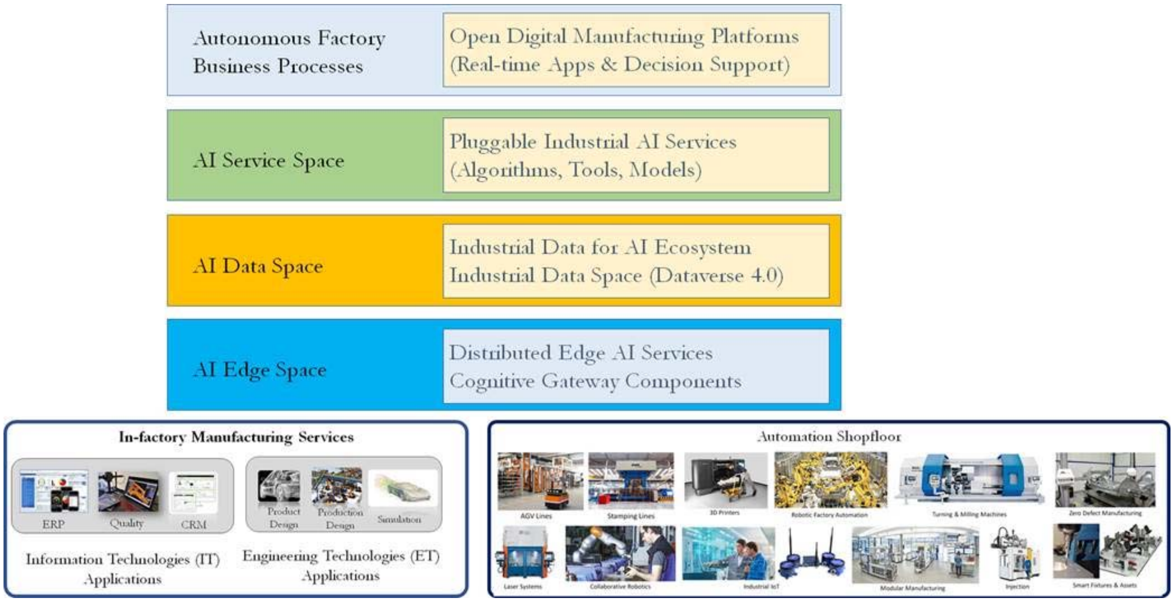


Figure 12 AI framework for SMI

Manufacturing is an extremely distributed environment characterized by very diverse and dynamic value chains that is being challenged by the increasing customisation of final products that call for increasingly collaborative and flexible production methods. Overall, this means that more autonomous decisions, yet with an overall coordination and common goals, are the aim of future manufacturing production environments. For this reason, intelligence should take different forms to fulfill very specific and high value missions for manufacturing operations. The requirements in the following sections are intended as technology enabler developments that will unveil AI across manufacturing products, assets, process, systems (factory) and service levels. AI will ultimately enable the implementation of new forms of decision and control loops that will enable more optimum, circular and sustainable design, commissioning, operation and after-sale service delivery in smart connected factory ecosystems.

Therefore, the new forms of AI should equally serve high-speed shopfloor event driven fast decision autonomous control and decision loops that will bring continuous manufacturing assets to their individual best performance, and autonomous collaborative multi-stage and multi-level decisions and control loops for cognitive optimum coordination of manufacturing execution tasks. Moreover, *deep analytic* control loops should be combined with other forms of control and decision loops so planning can become more dynamic and effective in addressing unexpected changes and seamlessly dealing with variation of conditions in global manufacturing and supply networks or events taking place on the shopfloor. Equally, humans should be always in the loop and assume a different role in terms of collaboration with AI based on the autonomy level set for the various processes. Human and AI (robotics & data-driven algorithms) need to improve the cooperation and distribution of tasks to make the best of both capacities (creativity, automation, speed…).

Thus, the overall use of resources and decisions can be balanced and reach the required level of flexibility. The implementation of such manufacturing operations calls for the implementation of four very distinctive types of smart manufacturing intelligence that reside and propagate across the layers defined in the previous Figure 12 and represented in Figure 13, namely Self-Adaptive Manufacturing Asset Intelligence, Cognitive Hybrid Twin Distributed Intelligence, Collaborative Assisted Intelligence and Human-centric Augmented Intelligence.

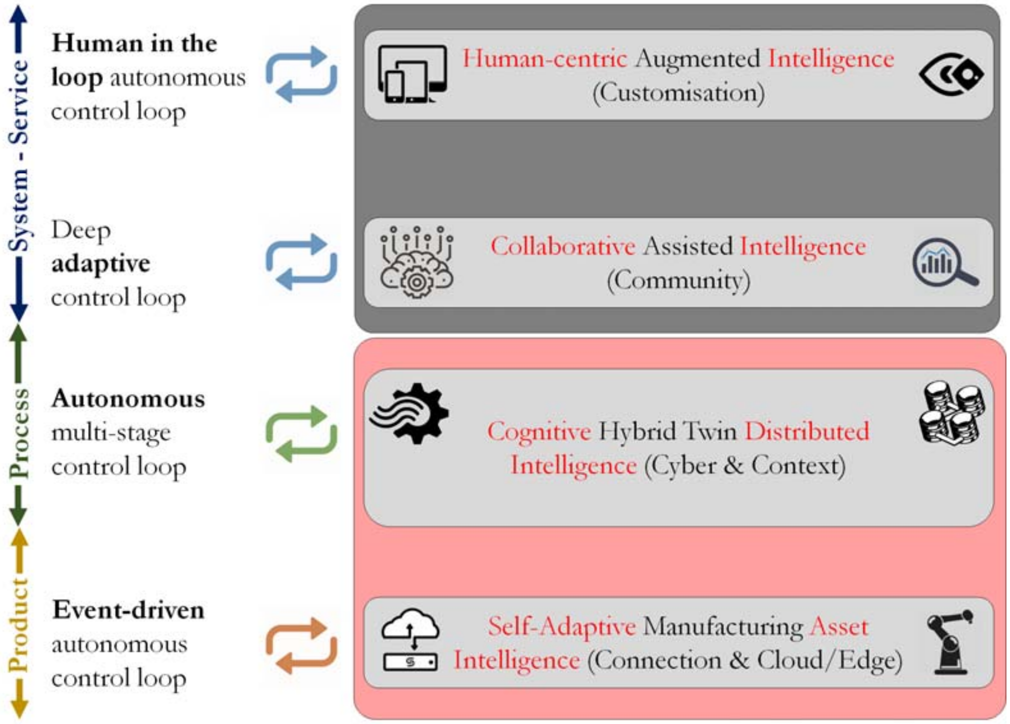


Figure 13 Implementation of AI-based manufacturing operations

Overall, the progress of smart manufacturing towards cost and time effective autonomous cognitive operations face three key challenges: (1) **COLLECTIVE** AI model generation across data assets and stakeholders in the data value chain, (2) human and AI **COLLABORATION** and (3) **CONCURRENT** and **COEXISTENT** optimization of management of cognitive service and product across lifecycle operations. There is a lack of trust and control on data usage, lack of understanding of AI reasoning from a human perspective as well as self-learning capacities and deployment of AI solutions are so far restricted to isolated processes, whereas cognitive factories demand that operations of a large number of such processes are concurrently and continuously optimized and adapted. The next sections discuss the challenges faced in leveraging the four types of manufacturing intelligence required to support new levels of autonomy (human-AI collaboration balance) in smart connected factory ecosystems.

4.3.1 AI in the Machines for Smart Manufacturing

Machine Learning (ML) is one of the AI areas that has progressed significantly. It is now a core technology for controlling the machines and for enabling smart and error-free operation. The following list focuses on the challenges for applying ML to Smart Industries:

Optimizing Production Chain. Utilize AI technology for optimizing production chains especially for high-variability products. For the single machine, the challenge is to integrate its operation into the overall optimization and to ensure individual operation, as well as overall optimized production. The challenge is how to address local optimization of manufacturing asset configuration with global and factory value chain manufacturing planning and organization.

Failure Detection. Utilizing AI for detecting failures (part quality, asset maintenance, supply chain, material, assembly, product, service) in the production chain and for proposing re-planning and re-scheduling actions is very attractive to the equipment manufacturing industry, but still has a very limited adoption. The main challenge, for ML, is that while considerable training data is available for correct operations, for failures, the amount of data is limited. While the general occurrence of a failure can be found using "Anomaly Detection" principles, the ability to correctly diagnose the failure and identify the root cause, is very limited. Especially for rare failure cases, data for ML might not be available. Challenges are whether such rare failures can be simulated and this data used for ML training. A novel area is to combine the artificially generated data with the usage labelling functions from human experts for diagnosis, root cause analysis, and the prediction of problems.

Operation & Maintenance. The vast amount of data collected online by sensors while machines are operating makes it possible to predict their behavior and evolution along their service life and, thus, identify the best moment for maintenance activities (in terms of minimizing costs, down times, disturbance to customers, etc.). This approach is starting to be followed in many industrial sectors under the term condition-based maintenance (CBM).

Situation Awareness. Optimization and fault detection in Smart manufacturing needs to have an accurate situation classification for their correct understanding. Ideally these situations are detected in real-time and might serve as input to classification algorithms detecting complex situations. Creation and management of these complex situation classification graphs needs to be managed. Moreover, explainable AI methods that support human intervention and understanding are critical for the final application of such advanced decision control loops.

Data Labeling Effort. Utilizing a data driven approach for training AI components in the smart manufacturing area requires a large range of datasets that are properly cleaned and labelled. The effort required for generating these training datasets is huge and needs to be minimized. Novel approaches (e.g. artificially generated data, labelling functions, data programming) need to be evaluated for reducing effort and cost.

Concept Drift. Trained AI models are achieving high accuracy when solving the given task for which they have been trained. In the course of the operation, the available data (labelled / unlabelled) is increasing and therefor the underlying statistical model that is detected by the AI might be changing.

Problem Learning. An important aspect for AI is to learn how to identify unusual situations and the circumstances, or values, where errors or faults occur. However, as these cases typically happen very infrequently in practice, the proper data to train ML models to anticipate those cases is usually not available.

Transfer and Multi-Task Learning. A trained model is valid for the situation where the training data was collected and labelled. Applying an already trained model to a new situation (e.g. a different machine or location in the world) typically results in inaccurate and imprecise results. Transfer learning is looking at how to modify a trained model so that it can be safely applied to a new set of input data.

Combining physical models and data-driven approaches: Physical models using various mathematical models can be used to describe the behavior of a machine and a production line. But due to increasing digitalization as well as the complexity of the machines, creating such mathematical models has become more and more complex and tedious. Expert level knowledge is needed on the mathematical methods as well as on the working of the machine. On the other hand, data-driven approaches have been successfully applied utilizing a different skill set and methods. The challenge in Smart Industry is now to combine those two approaches. How data-enrichment with physical and digital twin models of manufacturing equipment could be exploited is still a challenge if it needs to be brought to a suitable level of automation and human intervention in the generation of meaningful digital scenarios is to be reduced to only high added value tasks.

4.3.2 AI at the Edge for Smart Manufacturing

Edge computing is adding computational resource into the field. Those mini-data centers can provide computing resources (CPU, Storage, algorithms, AI) on demand. The following text explains the challenges for the Smart Industry area.

Computing Models for Smart Industry: Like cloud computing, edge computing needs the right computing model that a runtime system can use to manage the orchestration of resources for a task. Typical edge computing models are “Simple Task Assignments”, “Data Stream Processing” and “Service-oriented Computing”. For a factory, there are challenges in organizing the computing resources to fit to the structure of the production, to follow, automatically, changes in the production line, as well as to optimize the edge-cloud processing. Reinforcement learning might be applied to convert to an optimal resource assignment.

Offline Edge Training. Offline training of AI components means the components are trained with data collected from the target system, but then trained in a dedicated training task. The resulting model then needs to be deployed on the edge. This enables the use of the speed of the cloud and its data management capabilities to train the model. Afterwards the model can be deployed to the target system, ideally running on an edge component.

Online Edge Training. Online training means that ML components are trained while the system is running. This means: (a) training data needs to be generated during runtime (putting extra effort on humans or the automation of ground truth labelling), and (b) having the means to incrementally update the trained model. Online training has the advantage that the trained model is adapting according to changes in the behavior of the running system.

Edge Hardware for AI processing. Every day, new hardware that is optimized for AI processing, is becoming available. Edge clouds need to make this advanced hardware available to the deployed AI components. Edge hardware can include high-speed storage (e.g. NVMe), special co-processor for AI, special network functions for organizing the orchestration of edge computing, or especially encryption offloading technologies.

Edge Clouds. Edge nodes need to processes increasingly larger data sets. Those edge nodes need the virtualization and resource management functions of real clouds, adapted for the resources available at the edge. Furthermore, we need programming that enables developers to transparently utilize a cluster of edge nodes, as well as to utilize the large variants of edge resources that are becoming available.

Heavyweight and Lightweight Edges. Edge nodes can have a large variety of performance characteristics (e.g. dedicated application processors vs. re-used left over CPI cycles). Typically, heavyweight edges are dedicated to provide computing, storage and networking resources while lightweight edges are enabling AI as co-processes of existing nodes.

Cooperating Edge Nodes. Cooperating edge nodes exchange data and computation to jointly provide the needed service.

Resilient Edge Processing. As edge nodes are allocated close to the manufacturing processes or even the real-world systems, edge processing needs to be resilient against all kind of failures, e.g. partial failure of some nodes.

AI-based Data Ingestion. Today's IoT systems make IoT data available using pre-defined connectors. Future edge systems will dynamically explore IoT resources at the edge and make them available. Knowledge Extraction is an AI function that tries to understand incoming data and classify it according to the current knowledge ontology.

Ontology Matching: It is anticipated that that future hyper-connected Smart Industry will be operated based on agreed standards, however, the risk is that it will be nearly impossible to have always up-to-date standardized data models for every new machine introduced into the Smart Factory. Therefore, new ways of learning to integrate a new data model into the AI system controlling the Smart Factory need to be found. Ontology matching is one technique that needs to be mastered in the Smart Industry domain for rapid data and machine integrations.

4.3.3 AI in the Cloud for Smart Manufacturing

Due to extreme dynamism of the manufacturing context, use cases require flexible and efficient computing infrastructure, which spans from edge to cloud and back, creating a computing continuum, where AI is needed for resolving key challenges, like (1) how to take into account what, where and when data is collected and analyzed; (2) how to design services to respond to changes in application behavior or data variability; and (3) how to react to changes and trigger rules associated with the content of the data and models. The following list focuses on the challenges for AI in the Cloud for Smart Manufacturing Industry:

Intelligent Data to Cloud: Intelligent Streaming in the Cloud. Enabling intelligent processing of input data on the way to the Cloud. AI is used for deciding which data is relevant and how it should be cleaned to leave it properly stored for further processing. It also includes a dynamic (re)placement of the data processing along that path. This is important for the manufacturing domain since the data sources are usually unreliable, i.e. various types of inconsistency can appear in the data.

Intelligent Data preparation. Supporting the preparation of data for various processing methods (cloud). For example, using AI for creating missing training data (so called Artificial contrast data).

This is important for the manufacturing domain since the negative training examples are usually missing (e.g. for anomaly detection or out of control data).
AI Deployment in the Cloud - MLOps for Manufacturing. Machine learning operations (MLOps) is the use of machine learning models by development/operations (DevOps) teams. MLOps seeks to add discipline to the development and deployment of machine learning models by defining processes to make ML development more reliable and productive. This is important for the manufacturing domain since the deployment of AI applications can be constrained in different ways.
Complex (hybrid) processing pipelines. Resolving complex manufacturing problems requires orchestration of various AI methods, mainly composed of model-based and data-driven approaches. It requires complex workflows which can be dynamically changed based on the variability in data or application requirements. It also includes the orchestration of distributed processing elements in the case of decentralized scenarios. Such a hybrid processing is the basis for analysing and resolving previously unseen problems (so called unknown unknown situations), leading to realizing capabilities for cognitive processing.
Hybrid modelling and analytics. Data-driven models attempt to capture the system behaviour using only input and output data from the manufacturing line. Even though they are accurate for specific datasets, the models are highly dependent on the choice of training data, and their description of process physical phenomena is not explicit. First-principle-based models simulate the process with known mechanistic equations, but it is hard to model the complex behaviour of the line completely. AI will support an efficient, goal-oriented and dynamic integration of both approaches. For example, machine learning can be used for analysing the large data set, and the resulting data-driven models can be used to enhance first-principle models for the development of a hybrid model.
Hybrid Digital Twins. Beside the expansion of available data sources, the manufacturing domain is extensively using numerical models, especially for providing simulations of various difficult-to-measure parameters. To support an efficient integration, these structures (data, model) should be represented in a common format, leading to the concept of Hybrid Digital Twins that models the usual behaviour of the underlying manufacturing system/process/asset. The role of AI is to provide (background) knowledge for supporting semantic integration of heterogeneous concepts and enabling completely new services offered by Digital Twins, like providing data-driven anomaly detection on simulated data, resulting in new types of conclusions and corresponding proactive behaviour.

5 NON-TECHNICAL CHALLENGES FOR SMI IN DIGITAL EUROPE

A Data-oriented Digital Transformation for the Manufacturing Industry is not just technology innovation and take-up. Several non-technical challenges need to be considered and addressed, such as regulatory and legal challenges (§5.1), innovative revenue and business models (§5.2), new roles and professions and skills development programs (§5.3).

5.1 Regulatory and Legal Challenges

After a deep desk research and conjoint analysis of documents, six challenges for smart manufacturing for the policy, regulatory and certification framework were identified, as follows:

1. **'Servitisation' of manufacturing and the legal evolution of product as a service**
   The manufacturing industry is challenged by what has been termed 'smart factories', 'smart industry' or, simply put: 'servitisation' of the manufacturing process. Servitisation unveiled the role and impact of new actors in the manufacturing process while at the same time it blurred the distinction between the phases of 'manufacturing' and 'operation'. The focus has therefore shifted from the product to the ecosystem of actors and artefacts. For instance, in the car manufacturing industry, a manufacturer is no longer simply a manufacturer, much like a carrier is not merely an operator: their roles are intertwined and increasingly interdependent.[22] In a continuously evolving ecosystem of actors and cyber-physical artefacts, it is no longer clear-cut when the manufacturing process ends and when operation begins, for the two processes are intertwined. The challenges to the legal framework are significant and materialize especially in the domain of liability. Questions arise whether tort law can deal with the allocation of liability in situations where conduct is mediated through an autonomous system. The adequacy of product liability regimes based on a rigid distinction between product on the one hand and services on the other hand, is also challenged.

2. **Moving beyond deterministic approaches to certification**
   Smart manufacturing is marked by the use of autonomous systems. These systems do not follow a linear model of development; in a sense, they are never truly 'finished', their goals may evolve as they learn, and their behaviour may be assessed according to different legal frameworks and ethical standards at different points in time. Current approaches to certification and standards do not reflect this. These approaches are based upon an understanding of products, parts and components as "static" elements which do not change beyond certain predetermined limits. Efforts to integrate legal considerations into engineering have been limited to embedding strict deterministic design and certifying technical requirements into legal and regulatory instruments. The current way of standardization and certification can at best 'codify' the "desire that such systems be reliable"[23] but it is not sufficient to meet the new challenges. Safety regulations and assurance practices illustrate this. They assume that the full behaviour of the system is known at the design stage and, hence, that it can be assessed prior to system deployment. Safety assurances have therefore been integrated into the design processes, and compliance to safety standards has been sought during the system's test phases. Current safety standards and assurance approaches assume that the system, once deployed, will not learn or evolve. However, these standards and approaches do not extend to continuously evolving systems. Thanks to advances in AI, autonomous systems will have the potential to learn from their mistakes and those of all the systems they are connected to. This will infinitely enhance their abilities to operate safely. At the same time, AI brings along uncertainty about how the system will react to a particular future circumstance. Together, these developments significantly

---

[22] C Ducuing, I Emanuilov and O Dheu, 'The Emperor's New Clothes: a Roadmap for Conceptualising the 'New Vehicle', 2019 (unpublished).
[23] D Danks and AJ London, 'Regulating Autonomous Systems: Beyond Standards', *IEEE Intell. Syst.*, vol. 32, no. 1, pp. 88–91, Jan. 2017.

complicate the issue of safety assurance. Adaptive systems hence require adaptive certification and regulatory processes that radically differ from traditional legal approaches to certification. At the same time, the principle of adaptive regulatory processes seems to be at odds with legal principles such as legal certainty, which implies, among other things, that laws should be foreseeable, stable, non-retroactive, and consistent. Moreover because of its closely intertwined engineering and legal aspects, addressing this issue now requires a novel interdisciplinary approach.

3. **Supply chain legal accountability and traceability**
   In a smart manufacturing context, the determination of "who controls whom" and "who controls what" or, in other words, who is accountable and for what, becomes increasingly difficult. Various regulatory tools, such as contracts and the existing national liability regimes, have been put forward as possible solutions. However, contracts suffer from a major deficiency: they produce legal effects only between the parties. This is a strong limitation in complex, multi-actor environments, whether for liability allocation or law enforcement. With the exception of manufacturers' liability for defective products, which is partially harmonised, the existing liability regimes' fitness for purpose is also questionable since they are largely national and divided between the traditions of the common and the civil law. This fragmented legal landscape makes it difficult to allocate the duties and, hence, the liability in a smart manufacturing environment which extends beyond the physical boundaries of a factory. Furthermore, the emergence of new legal frameworks for the digital environment exacerbate these problems. The dependence of smart manufacturing on cyber-physical systems brings an additional set of legal questions to the fore regarding the (cyber)security, safety and compliance of these systems and operations with international human rights law and international regulatory law. As a global actor, the EU has already taken several harmonisation measures to tackle some of these challenges, such as the Directive on Critical Infrastructure Protection, the Network and Information Security Directive, the EU Cybersecurity Act, the Seveso regime and the General Data Protection Regulation. However, many of these efforts have limited impact as sometimes they leave significant room for manoeuvre for Member States or do not address emerging concerns, such as the convergence of safety and security requirements or the need for mandatory human rights due diligence for companies in the 'digital supply chains' for compliance with international human rights standards.[24] These legal challenges have not only an external, actor-based, but also an internal, system-level dimension. The legal accountability framework should be advanced. In this context, the datasets and processes that lead to a certain AI decision should be documented with the aim to ensure traceability. This component encompasses both *ex ante* allocation of duties and imposition of due diligence obligations and *ex post* distribution of liability in collaborative manufacturing supply chains for continued compliance with international and EU human rights law.

4. **Uncertain legal status of the production, use and sharing of data**
   Smart manufacturing is marked by a multiplicity of actors in the value chain providing data which is necessary to make real-time decisions across the value chain. In such a scenario data is an economic asset, access to and re-use of which should be encouraged. The economic value of data as the fuel for the digital economy questions their legal status, particularly in light of recent 'ownership' claims by some stakeholders. Even in jurisdictions where statutory law provides for, or is silent on, ownership rights in intangibles, questions arise as to the "materiality" of data, given their specific features.
   Additionally, smart manufacturing builds upon digitization of physical goods and processes, illustrated by the creation of digital assets such as digital twins. This blurs the lines of what is or should be a legal object – data or (some) digital assets, and raises additional challenges for the legal framework. Beyond the realm of property law, data is often affected by a patchwork of

---

[24] L McGregor, D Murray and V Ng, "International Human Rights Law as a Framework for Algorithmic Accountability" (2019) 68 International and Comparative Law Quarterly 309. See also Human Rights, Big Data and Technology Project and Essex Business and Human Rights Project, 'Submission by the Human Rights, Big Data and Technology Project ("HRBDT") and the Essex Business and Human Rights Project ("EBHR") to the UN Working Group on Business and Human Rights ("UNWG") for the Consultation Process to Inform Its 2018 Report to the UN General Assembly' (2018), <https://www.ohchr.org/Documents/Issues/Business/WGSubmissions/2018/Essex.pdf> accessed 4 March 2019.

rights, which sometimes conflict with one another, as a single data (operation) may simultaneously be impacted by several legal regimes (e.g. data protection law, intellectual property rights, contract law, trade secrets, various obligations to give access to certain data, etc.). Often, the legal regimes do not tackle data as such, but data on the ground of some other feature, e.g. the protection of privacy and self-determination of a data subject (data protection law), the investment and work of the creator of a database (*sui generis* legal protection on databases), etc. As a result, the determination of who is entitled to perform what activities on what data becomes a difficult legal question in the context of "big data" and results in legal uncertainty.[25]

5. **Exploring the role of self-regulation as a complementary regulatory instrument**.

In a data-driven Manufacturing Industry, the possible role of self-regulation at company and eco-system level should be further explored and supported by governments and institutions, especially in relation to the deployment and uptake of AI-empowered systems. Efforts should be directed to investigate the expected effects of a shift from the vision of mere legal compliance obligations towards incorporating the principles and KPIs of trustworthy AI in self-regulatory tools[26]. This might entail their introduction, on the one hand, as part of a broader **Corporate Social Responsibility** (CSR) or deontology charter, and, on the other hand, within sector-specific Codes of Conducts.

The possible benefits of the "soft law"[27], its relationship with the traditional legal instruments, at national and EU level, as well as its role in a landscape of increasingly dynamic cross-fertilization of regulations and technology, should be further investigated, even more so in rapidly developing fields like AI-empowered SMI. Soft law could contribute to paving the way for ensuring that AI and data innovation are deployed for better social outcomes, and providing important safeguards on issues like transparency and accountability. Being flexible and able to be quickly adapted to future technological progress, soft law could overcome the risk of the lack of timely alignment of the current legislative system, which develops at a much slower pace than the AI and data-driven technology and its applications.

As regards the role of soft law at eco-system level in the Smart Manufacturing environment, the possible advantages of the production of realistic and workable ethical Codes of Conduct for each domain should be explored and experimented with. Such regulatory sources are coherent with the legislative support for the self-regulation and accountability instruments (e.g. GDPR, art. 40). The Codes should be created by bodies representing each specific sector, such as trade associations or Digital Innovation Hubs. They are expected to bring a set of advantages, being capable of offering guidance and addressing in meaningful, flexible and practical ways the immediate and realistic issues and ethical challenges of AI breakthroughs in each of the SMI Grand Challenges, going beyond current gaps in the European and national regulatory framework. They can provide a tailored description of what is the most appropriate, legal and ethical array of behaviours in each SMI domain, concretely operating as a rulebook for developers, uptakers and other relevant stakeholders who design and implement AI systems. Codes can provide a degree of co-regulation and a more granular ethical guidance for the specific AI activities within each domain, providing practical solutions to any problems and challenges identified in it. This would result in an increase of confidence and legal certainty and would earn the trust and confidence of individuals.

In the data protection sector, the European Data Protection Board set a number of requirements for the Codes of Conducts relevant to GDPR[28]: they could likewise be taken into consideration and adapted for the wider concept, depicted above, of trustworthy AI Codes of Conducts in the SMI. These Codes, enabling trustworthy AI compliant business models upholding EU values,

---

[25] C. Ducuing, D4.5 – Legal aspect for smart contract adoption, In2Dreams (2018).

[26] The principles and requirements for designing, deploying and adopting Trustworthy AI systems have been outlined by the High-Level Expert Group on AI, set up by the European Commission, in the "Ethics Guidelines for Trustworthy AI" (8 April 2018): https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

[27] Instruments, such as codes of conduct, guidelines and roadmaps, not legally binding, or whose binding force is to some extent "weaker" than that of traditional or hard law.

[28] EDPS, Guidelines 1/2019 on Codes of Conduct and Monitoring Bodies under Regulation 2016/679, adopted on 12 February 2019. The requirements includes, for instance, representativeness, consultation, defined scope, compliance with the applicable law, review mechanisms, scrutiny, effective oversight and monitoring mechanisms and enforcement measures.

would facilitate the respect of the European approach to AI, relying on such values. Such an approach is perceived as an enabler of the European global leadership in AI, as proven for instance by the EC's Communication "AI for Europe"[29] (25 April 2018) and "Building Trust in Human-Centric AI"[30] (8 April 2019)[31].

Soft law and Code of Conducts could, therefore, contribute to fostering data exchange and unlocking AI potential, thereby helping EU SMI to materialize as an "AI-multiplier" effect by promoting an ethically-driven approach to AI development at the global level, whilst nurturing social acceptance.

6. **Safeguarding human self-determination and fostering human continuous empowerment**
   The disruptive automation technologies, including above all human-machine interaction, AI and robotics, pave the way for unprecedented transformation of our Manufacturing Industry and economic growth: this raises the challenges for the legal and ethical framework to ensure that such data-driven advances respect individual rights in each of the three Grand Scenarios identified in the SMI subgroup Discussion Paper released in 2018[32] (Smart Factory, Smart Product and Smart Supply Chain scenario). As a minimum, human dignity and agency must be preserved, in terms of ensuring that the individuals interacting with AI-powered/automated systems retain control and self-determination over themselves, including in the workplace. Novel practices, such as dynamic user consent and sticky policies, should be investigated as instruments for this purpose. Human beings must be fully informed of any form of interaction with such systems, keeping the right to decide whether to be subject or not to their direct or indirect decision making. The right to opt out and withdraw must be ensured without any consequent discrimination based on it.
   As an evolution, in line with the "sharing the wealth" paradigm[33], a pathway towards pursuing continuous human empowerment must be fostered to the maximum extent. Above all, this concerns the Autonomous Factories persona related to the Smart Factory Grand Scenarios and its topic Workplace Human-Machine interaction. Employees should feel empowered and augmented, not threatened by AI technologies. Vulnerable groups, such as workers with disabilities, must not only be safeguarded but even fully valorised to effectively let them, as individuals as well as groups, increase their degree of autonomy, self-determination and consequent satisfaction, unlocking their potentialities.

Further specific interventions at policy and regulatory level are needed to ensure that Europe's AI talent base and workforce are strengthened: skills development and skills-oriented action (including retraining) should be inclusive, not discriminatory and directed to human empowerment, whilst avoiding the risk of exacerbating power asymmetries, algorithmic bias, digital skills divide, as well as distress and feelings of frustration.

The effect of technology on employees', customers' and other citizens' well-being in the manufacturing environment must be addressed in view of long-term inclusive digital "prosperity", setting the ground for a broad level of acceptance of AI technologies. This recommendation is aligned with the EC's strategy[34], likewise with the EFFRA[35] 2020 Roadmap and its Vision for a Manufacturing Partnership in Horizon Europe 2021-2027[36], which are directed to foster the

---

[29] COM(2018) 237 final

[30] COM(2019) 168 final

[31] The same vision is also common to key initiatives, promoted by the EC, such as the Strategic Research, Innovation and Deployment Agenda for an AI PPP - A focal point for collaboration on Artificial Intelligence, Data and Robotics (consultation release published on June 2019), which is a joint effort of BDVA and EuRobotics: Zillner, S., Bisset, D., García Robles, A., Hahn, T., Lafrenz, R., Liepert, B., and Curry, E. (eds) (2019) "Strategic Research, Innovation and Deployment Agenda for an AI PPP: A focal point for collaboration on Artificial Intelligence, Data and Robotics", Brussels. BDVA – euRobotics

[32] BDVA, Big Data Challenges in Smart Manufacturing, A Discussion Paper on Big Data challenges for BDVA and EFFRA Research & Innovation roadmaps alignment, 2018

[33] O. Tene and J. Polonetsky, Big Data for All: Privacy and User Control in the Age of Analytics, Northwestern Journal of Technology and Intellectual Property, Vol. 11, 2013.

[34] For instance: COM(2019) 168 final "COM "Building Trust in Human-Centric AI".

[35] European Factories of the Future Research Association: http://www.effra.eu/

[36] https://www.effra.eu/sites/default/files/190312_effra_roadmapmanufacturingppp_eversion.pdf

deployment, use and consequent perception of such technologies as useful and comfortable tools to complement and enhance human capabilities, rather than feeding fears that they will replace human workers throughout the SMI domain. In this direction, the synergy with initiatives like the IEEE Global Initiative on the Ethics of Autonomous and Intelligent Systems[37] should be pursued: its standardization projects for prioritizing Human Well-being in A/IS (IEEE P7000 series[38]), addressing specific issues at the intersection of technological and ethical considerations, should be carefully monitored.

In this perspective, the concept of Collaborative Intelligence (CI), coined by James Wilson and Paul Daugherty[39], sounds promising. It aims at harnessing the capabilities of machines for enhancing human capabilities in a way that prioritizes human flourishing, in particular throughout two bidirectional collaborative interaction channels empathizing Human Factor's role in AI and A/IS technologies in the perspective of 2025 EU Connected Factories scenarios: "Humans Assisting Machines" and "Machines Assisting Humans".

The concept of CI is aligned with the vision and roadmap of the Council on Extended Intelligence[40] (CXI), created by the IEEE Global Initiative in conjunction with MIT Media Lab: the further developments and insights of both should therefore be assessed and taken into account in further research initiatives, thereby moving towards a global perspective of AI ethics strategies relying on data sharing, fostered by the EC[41].

---

[37] https://standards.ieee.org/industry-connections/ec/autonomous-systems.html

[38] Some examples:

- IEEE P7003™ - Algorithmic Bias Consideration
- IEEE P7006™ - Standard for Personal Data Artificial Intelligence (AI) Agent
- IEEE P7007™ - Ontological Standard for Ethically Driven Robotics and Automation Systems
- IEEE P7010™ - Wellbeing Metrics Standard for Ethical Artificial Intelligence and Autonomous Systems

[39] *"Collaborative Intelligence: Humans and AI Are Joining Forces"*, James Wilson and Paul Daugherty, Harvard Business Review: https://hbr.org/2018/07/collaborative-intelligence-humans-and-ai-are-joining-forces

[40] https://globalcxi.org/

[41] In the COM(2019) 168 final "COM "Building Trust in Human-Centric AI", for instance: the EC underlines that it "will continue its efforts to bring the Union's approach to the global stage and build a consensus on a human-centric AI"

5.2 Innovation and new Business Models Challenges

The digital transformation of the manufacturing industry in Europe takes place in an increasingly competitive global manufacturing market. Thanks to investments – accelerated by the use of dedicated EU funds – in resource efficiency, new technologies, skills, and access to finance, digitization has erupted as an opportunity for boosting the competitiveness and sustainable growth of the European manufacturing industry.

The sensorization and connection of physical products, resulting in the collection and analysis of large quantities of data, are creating big expectations as the technical possibilities of knowledge and value generation for the manufacturing industry. However, even if the industrial equipment is digitally enhanced, the product by itself is not able to deliver the increase in revenue expected from these technologies. The prerequisite for revenue expansion, in fact, is a **change in the business models**, which is reflected in the mechanisms of value creation, value delivery, and value appropriation adopted by manufacturing companies in sectorial value chains. Along these lines, a number of data-driven business model patterns are unleashing transformative forces at sectorial level:

- **Outcome-based model** (i.e., delivery mechanism in which the customer is charged by usage, or outcome, instead of for product ownership) – Examples: 'Power by the Hour' by Rolls-Royce (engine and accessory replacement service offered on a fixed-cost-per-flying-hour basis), Kaeser (shift from selling compressors to selling compressed air per cubic-meter).
- **Extended product** (i.e., tangible core product combined with value added data-driven services available throughout its lifecycle) – Example: MyJohnDeere (product-service-system coupling tractors with advanced analytics for performance optimization and predictive maintenance).
- **Re-manufacturing** (i.e., end-of-life treatment dismantling the product, restoring and replacing components and testing the individual parts and whole product to ensure that it is within its origin design specifications) – Example: Hitachi Remanufacturing Centre (supply of re-manufactured pumps, swing and travel motors for excavators).
- **De-manufacturing** (i.e., breakdown of a product into its individual parts with the goal of maximizing reuse and recycling opportunities) – Example: Airbus (85%-95% of aircraft components recycled, reused, or otherwise recovered).
- **Mass customization** (i.e., combination of flexibility and personalization that are peculiar to bespoke manufacturing with the low unit costs associated with mass production) – Example: Lanieri (production of tailored, made-to-measure men's clothing, made with 100% Made in Italy fabrics).
- **Virtual factory** (i.e., orchestration of a community of micro-factories, all linked by an electronic network that enables them to operate proximity production when receiving an order) – Examples: Local Motors (recourse to multiple micro-factories for short-run production of motor vehicles), Shapeways (digital manufacturing platform for 3D-printed consumer goods).
- **Crowdsourced innovation** (i.e., exploitation of collective intelligence with the aim of gathering valuable design ideas and technical insights to be turned into real products) – Example: P&G Connect + Develop (50% of new product/process innovations originated from outside the organization through crowdsourcing challenges).
- **Analytics-as-a-service** (i.e., on-demand access to a one-stop-shop analytics engine supporting shop floor intelligence) – Example: AVL data intelligence services (support to the complete vehicle development process from concept phase to after-sales management as part of a life-long partnership).
- **Open data platform** (i.e., data hubs meant to integrate disparate data sources, thus enabling information to become shareable and actionable) – Examples: Skywise (Airbus- Palantir initiative to put every aviation stakeholder in the position to enhance decision making on the basis of operational data collected from thousands of aircrafts), Predix (GE platform providing industrial-scale analytics for asset performance management and operations optimization based on IoT data from industrial machines).
- **Dynamic pricing** (i.e., SLAs, pricing plans, payment terms, and other contractual terms may be customer-specific or based on real-time market factors) – Example: eMachineshop (instant pricing for on-demand manufacturing services).

Two common threads run through aforementioned approaches: they are the trajectories with the farthest-reaching implications.

Firstly, a new wave of servitization is propagating in the manufacturing industry. In this context, besides smart goods and smart production environments, the smart servitization concept is gaining traction. Servitization is a term that has been used in the manufacturing sector for a long time, while pay-per-use is starting to spread from just a few industries, like electricity or phones, to more traditional sectors. IoT and Industry 4.0 paradigms are setting a technological basis for these new business model developments, as testified by new terms such as PSS (Product Service Systems), uptime guaranteed, FCMs (Flexible Consumption Models), also known as “as-a-Service” or XaaS models. When it comes to economics, equipment OEMs usually have between 8 to 15 percent margin on new product sales. Aftermarket service can generate up to eight times that revenue in profit over the lifecycle of the equipment. This is why service is so important as a profit driver. It is not unusual for companies to generate 18 to 20 percent of their total profit from service operations even though service revenue may only account for 5 percent of total revenue. Taking this a step further, equipment with small sales margins can benefit greatly from offering the asset as a service[42].

Secondly, circular economy represents a fundamental avenue for minimizing waste and making the most out of resources, in contrast to the ingrained linear, take-make-waste production process. Traditional business models (BMs) prioritize the sale of products as they are the main source of revenue, penalising environmental aspects. The new BMs, where the income of the manufacturer of the asset depends on the use of it, favour the circular economy (CE) paradigm. In a circular economy the value of products and materials is maintained for as long as possible and waste and resource use are minimised[43]. Digital technologies are enabling this transition as they provide continuous knowledge of the usage of the assets. In the pay per use business models the provider is responsible for all lifecycle costs and supported by IoT, Big data and machine learning, it can achieve the following objectives of the CE [44]:

- **Design products for CE**: based on collected usage data they can be redesigned to be maintained, upgraded, disassembled, and recycled in an easier way
- **Minimize operating costs** (for instance, by increasing resource efficiency): extending the product lifespan, and collecting back products to allow multiple lifecycles
- **Monitor the product** condition, status, location, and usage for proper use avoiding wear and tear
- **Introduce sharing BMs**: depending on use and need of production of different manufacturers. Increases the product utilization, improvement in resource efficiency
- **Improve the provision of technical support and other services**, such as repair, assistance, spare parts management, etc. so that the lifespan of the product may be extended. Furthermore, companies may provide to their customers personalized advice with the aim of optimizing the usage phase
- **Reduction of the product consumables** (e.g., energy) during the usage phase, thus increasing resource efficiency.
- **End-of-life collection**, refurbishment, remanufacturing and recycling activities in a proper way, collection activities when products reach end-of-life may be optimized, since companies know each product location in real time.

Ground-level innovation in relation to servitization and the circular economy is being driven by large companies who have the resources to pilot business models based on leasing, product performance, remanufacture, and extended lifecycle thinking, and carry the risks associated with providing services. However, such efforts also rely on SME involvement, and in some respects that is proving to be a

---

[42] https://www.ifsworld.com/corp/sitecore/media-library/assets/2018/05/30/industrial-servitization-and-field-service/

[43] https://ec.europa.eu/growth/industry/sustainability_en

[44] Exploring How Usage-Focused Business Models Enable Circular Economy through Digital Technologies. Gianmarco Bressanelli et al. 2018

challenge. A recent survey of nearly 300 small businesses across England, France and Belgium found almost 50 per cent had not heard of the circular economy.[45]

Changing the business model is not a trivial exercise and industrial companies struggle to create value from these technologies, due to the scope and complexity of the changes that are needed. Flexible Consumption models[46] (as-a service, XaaS) offer the possibility of purchasing the products as a service with payment models for the function delivered instead of the product itself. This is a big mind-set change in the operating model for the industry. The machine supplier has to change from a sequential operation model **centred on the product** to a data-driven interconnected model with short planning cycles that is **client centred**. The change is disruptive for the traditional manufacturing sector as it has a long reach in the operations and the way they deliver value to the market. The IFS research study of 200 executives to determine how industrial companies are progressing towards servitization – the expansion of manufacturing and product-focused companies into value-added service after the sale – reveals that companies are still in the very early stages of business model transformation through digitalization. Only 4% of respondents reported full servitization, i.e., selling products on a subscription rather than as a discrete item, through power-by-the-hour, fee-for-usage or revenue sharing agreements.

Manufacturers need to work on pilot cases so they can gather insights and best practices to develop their value proposition, revenue model, system characteristics and competitive advantage sources. In a context of hybrid product and service organizations, one of the biggest questions is how to place service within the organization and how to manage the hybrid service and offer avoiding the risk of cannibalization. According to research by Accenture[47], 81% of executives say they expect to manage multiple operating models in parallel in the future. Yet many struggle because they do not fully align their operating models with these new disruptive business models (digital operating models can improve incremental revenue growth by 10-20%, according to the same Accenture research). Transforming the business model to a product-as-a-service through IoT impacts the whole organization and companies face a range of considerations and strategic questions pertaining to customers, value proposition, business case, product development, distribution, IT and/or partners.

The following innovation paths[48] can be outlined from the analyzed literature:

- Capitalize technological innovation through developing flexible organizational structures and combining product, process and service innovation.
- Develop specific tools and methodologies for new business model design, under the new industrial digitalization paradigm.
- Invest in business model experimentation, embracing collaborative academic–practitioner initiatives to inform and support business model innovation over sustained periods of time, and linking science and technology to a wider population and encouraging closer cooperation between private and public sectors.
- Facilitate economic development through entrepreneurship and the innovation potential of small and medium-sized firms, thereby making business model education and the supporting tools accessible to firms; to equip for innovation potential, including ‘intrapreneurship’.
- Understand technological enablers of new opportunities such as digital twins, robotics, artificial intelligence and 3D printing to name a few – within their benefit, use case and holistic context.
- Diversify and tap into new revenue sources, leveraging new ecosystems, and (connected) data, to thrive and in some cases survive.
- Study and overcome barriers to B2B data sharing[49] (e.g., scarce incentives; lack of trust among parties; confidentiality issues and IP protection; uncertainty about safety, security, and liability conditions; difficult contractual enforcement; unequal bargaining power; uncertainty about valuing data; economic costs of sharing data).

---

[45] http://www.sandbirch.com/10-things-you-need-to-know-about-the-circular-economy/

[46] https://www2.deloitte.com/content/dam/insights/us/articles/4007_shift-to-flexible-consumption/4007_shift-to-flexible-consumption.pdf

[47] https://www.accenture.com/_acnmedia/PDF-67/Accenture-Strategy-Adapt-to-Survive-POV.pdf#zoom=50

[48] The business Model Navigator. Oliver Grassmann, Karolin Frankenberger & Michaela Csik 2014.

[49] https://publications.europa.eu/en/publication-detail/-/publication/8b8776ff-4834-11e8-be1d-01aa75ed71a1/language-en

5.3 Industry 4.0 and Data-oriented Digital Skills

5.3.1 Skills Competitiveness

Labour force preparedness is a key pillar for companies´ competitiveness and societal well-being. Employees with advanced and modern digital skills provide value to their employers by enabling productivity, product customization, short time-to-market, etc. Nowadays, the digital skills demanded by industries, which are willing to differentiate themselves in a global competitive market, surpasses the existing skilled people stock. Moreover, digital technologies advance at a pace that has not been seen before. Thus, industries need to modernize themselves and invest in human capital keeping a continuous observatory of their market´s digitalization.

Within the context of today´s competitive world, the young labour force needs to be digitally skilled and open to geographic mobility around the globe. To this end, EU international degrees and masters from reputed universities are essential building blocks. Besides, employees need to upgrade their skills in order to get higher quality jobs and to keep adding value to their employees. Cross-generational teams can be a practical option to learn on the job from young engineers.

Moreover, although governments at local, regional and national levels are fighting against non-digital unemployment in order to achieve unemployed inclusion into the modern working force through digitalization, still more efforts are needed in this direction. Hence, support for public and private stakeholders in their investment in the labour force’s digital skills is “a must” to progress towards an EU harmonized skills convergence. In this regard, three main challenges are described in the following section: skills and competence frameworks; competence development; competence recognition. For each of them, an outline of state-of-the-art results, which can be leveraged as a starting point for future strategic actions, as well as current gaps and avenues for further developments, will be presented.

5.3.2 Skills and Competence Frameworks

Developing a better understanding of which skills are needed (and will be needed in future) is important for individuals, companies, policy makers and educational institutions. Moreover, the creation of common frameworks for skills and competences is a prerequisite for effective skill matching and development planning with respect to Industry 4.0 and data science in manufacturing.

The Industry 4.0 skill catalogue

The Industry 4.0 skills catalogue developed by Politecnico di Milano[50], with the contribution of the BEinCPPS project, covers 100+ technical and managerial skills pertaining to six Industry 4.0 areas, which have been identified as crucial for the design, implementation and management of future manufacturing business models, processes and systems in the Industry 4.0 scenario.

FACTORIES OF THE FUTURE COMPETENCE AREAS

- *Industry 4.0 Strategy*: It refers to the formulation and implementation of the Industry 4.0 strategy, which is critical to organisational success, requiring a deep understanding of the industry and the implications of emerging technologies, societal and environmental trends for the wider business environment, as well as a high level of leadership and stakeholder management.
- *Smart Operations*: It refers to the competences for the designing, planning, organisation, direction, execution and improvement of industrial processes and systems exploiting Industry 4.0 technologies and data, to ensure that competitive and sustainability priorities are achieved.
- *Smart Supply Chain*: It refers to the understanding of supply chain digital technologies and the ability to use those technologies and data from multiple sources with the goal of horizontally integrating and dynamically reconfiguring resources and processes, to improve

---

[50] Pinzone, M., Fantini, P., Perini, S., Garavaglia, S., Taisch, M., & Miragliotta, G. (2017, September). Jobs and skills in Industry 4.0: an exploratory research. In IFIP International Conference on Advances in Production Management Systems (pp. 282-288). Springer, Cham.

the effectiveness and efficiency of the supply chain and to respond to the quickly changing environment.
- *Smart Product-Service System*: It refers to the competences needed for the design of "smart" and connected products, together with the services linked to them. Moreover, it covers novel competences related to the design and engineering process, such as the use of digital graphics, simulation and virtual and / or augmented reality.

TRANSVERSAL AREAS

- *Industry 4.0 infrastructure*: It refers to different Industry 4.0 technological areas (cyber-physical systems, industrial IoT, cloud computing, HMI, etc.) to develop and implement an Industry 4.0 architecture – with both hardware and software components - that is oriented towards the creation of value.
- *Big Data*: It refers to the capacity to design, develop and implement new systems and applications for data collection, analysis and management; the use of statistical techniques and analytics on available data to deliver novel insights and discover new relations for decision-making.

With respect to Big Data related competences, three EU projects have been working actively on initiatives to analyze data science competences: i) the EDISON project[51] ii) the EDSA Project[52] and iii) the BDVe Project[53]. In addition, the European ICT Professional Role Profiles version 2 has been officially published by CEN on 2018-08-29. In the followings some of their results are reviewed.

The EDISON project

The EDISON project defined the EDISON Data Science Framework (EDSF) which is composed of the following four components:

- Data Science Competence Framework (CF-DS), the definition of Competences for Data Science according to the e-CF 3.0
- Data Science Body of Knowledge (DS-BoK), for each competence group, the identification of knowledge areas and knowledge units.
- Data Science Model Curriculum (MC-DS), the definition of learning outcomes for each competence, and the definition of learning units.
- Data Science Professional profiles and occupations taxonomy (DSPP), a listing of 22 professional profiles in data science with their knowledge areas and units.

Edison's Data Science Competence Framework (CF-DS) provides the competences for data science professionals. Five groups of competences are defined:

- Data Science Analytics
- Data Science Engineering
- Domain Knowledge and Expertise
- Data Management
- Research Methods

The European ICT Professional Role Profiles version 2

The European ICT Professional Role Profiles make a key contribution to increasing transparency and convergence of the European ICT Skills landscape. It incorporates the competences of the European e-Competence Framework (e-CF, EN 16234-1). The 30 ICT Professional Role Profiles provide a generic set of typical roles performed by ICT Professionals in any organisation, covering the full ICT business process. Seven new profiles have been added in the second version including the role of data scientist.

The data scientist mission is defined as: (1) Finds, manages and merges multiple data sources and ensures consistency of datasets; (2) Identifies the mathematical models, selects and optimises the algorithms to deliver business value through insights; (3) Communicates patterns and recommends

---

[51] http://edison-project.eu/
[52] http://edsa-project.eu/
[53] https://www.big-data-value.eu/

ways of applying data. In summary it is stated that a data scientist leads the process of applying data analytics and delivers insights from data by optimizing the analytics process and presenting visual data representations. Overall, from an industrial perspective, the following summarizes the job descriptions that are usually sought in the data science domain:

- Data science job title 1: **data engineer** (a person whose primary job responsibility involves preparing data for analytical or operational uses).
- Data science job title 2: **data scientist** (a person employed to analyse and interpret complex digital data, such as the usage statistics of a website, especially to assist a business in its decision-making)
- Data science job title 3: **data architect** (a person whose primary job responsibility involves defining how the data will be stored, consumed, integrated and managed by different data entities and IT systems).

The results described above represent a key starting point toward a common, harmonized EU reference competence framework, aligned with the European Qualification Framework. Such a competence framework could be exploited to build and implement novel approaches/methods/tools aimed at supporting manufacturing stakeholders in competence assessment, HR planning, qualifications and certifications etc. Moreover, due to the increasing importance of Artificial Intelligence applications in manufacturing, it must be highlighted that further collaborative work should be promoted to define those competences necessary to design and operate Artificial Intelligence-based systems.

5.3.3 Competence Development

Big Data technologies will penetrate more and more into manufacturing business models and industrial processes. They will become more mature and widespread. Indeed, it will be necessary to diffuse the related competences at all levels of manufacturing organizations and promote a lifelong learning approach to stay updated with the future evolution of the technology. Two inter-connected pathways are needed to provide manufacturing with the necessary competences to fully reap the benefits of big data technologies: on the one hand, traditional manufacturing roles must evolve and acquire new Industry 4.0 and data-related competences; on the other hand, new roles specialized in handling big data are emerging but need to develop expertise in the manufacturing domain to be effective in this context.

Nowadays, there are lots of formal and non-formal learning programmes on data science. However, the main emphasis for the next years will have to be put on the possibilities of “re-training”, so that those experts who do not come from the IT domain can acquire the data science skills needed to work with (or even challenge) AI experts. Besides, data science professionals are rarely effective if they don’t have a good knowledge of the manufacturing domain, its requirements and culture. Therefore, they will have to acquire manufacturing-related skills in order to work in the sector. In this way the basis to bridge the gap between manufacturing and data science will be settled. Additionally, it is crucial to attract new generations of students into data science and manufacturing, and to provide them with the necessary competences to fulfil those roles that will be in need by the industry in future. In this respect, creating EU international degrees and masters from reputed universities are essential building blocks, as well as the modernization of technical & *vocational education and training programmes. To this end, collaborative schemes based on industry-research-education-government alliances must be fostered at the EU, national and local level.*

As an example, the EDSA project proposed a curriculum for data science. That curriculum was based upon what the EDSA consortium identified as core data science knowledge rather than the skills that might be needed for a particular job in data science. This curriculum was validated through different surveys. The EDSA curriculum consists of 15 core data science topics. Each of these topics has learning objectives, descriptions as well as resources and materials which were also produced as part of the EDSA project. The 15 topics that make up the core EDSA curriculum were divided up into four stages: Foundations, Storage and Processing, Analysis, and Interpretation and Use. Finally, novel methods and delivering mechanisms within new “cyber-physical” learning spaces (e.g. didactic factories, hackathons, etc.) must be explored and exploited to develop data science competences, and to bring them closer to industrial practice, in order to overcome the pitfalls of conventional education and training that are often disconnected from industry.

5.3.4 Competence Recognition

Documenting skills and qualifications in more transparent and understandable ways is important to facilitate the match between data science skills demand and supply in manufacturing, and beyond it. The shift from paper-based certificates to digitally-signed credentials represents a step in this direction. In this context, Digital Badges that comply with the Open Badge standard can be included in curricula, uploaded to platforms like LinkedIn, and shared on social media. In order to define both the types and requirements of the BDVe Data Science Badge program, BDVe developers have focused on EDISON's Data Science Competence Framework (CF-DS) and Data Science Model Curriculum (MC-DS). With the only variation being the change of the name of Badge 5 from Data Science Research Method and Project Management to Data Science Research Methods. In the BDVe project a framework to validate skills of data science has been developed. As an example the required skills for the data science analytics badge are listed in Table 1. A pilot of the academic level of the Data Science Analytics Badge is currently underway. Data science badges could find applications not only in formal education, in the sense that they can complement and enrich traditional credentials, but also in informal and non-formal education, by certifying competences achieved in all learning environments. Therefore, future efforts will be needed to scale up current initiatives.

Figure 14 Badge 1: Data Science Analytics

Table 1 Skills required for BDVe's Data Science Analytics Badges

Required Skills:		
Basic Level	**Intermediate Level**	**Expert Level**
Choose and execute existing data analytics and predictive analytics tools.	Identify existing requirements and develop predictive analysis tools.	Design and evaluate predictive analysis tools to discover new relations.
Identify existing requirements and develop predictive analysis tools.	Select most appropriate statistical techniques and model available data to deliver insights	Assess and optimize organization processes using statistical techniques.
Design and evaluate predictive analysis tools to discover new relations.	Analyze available data sources and develop tool that work with complex datasets.	Assess, adapt, and combine data sources to improve analytics.
Name and use basic performance assessment metrics and tools.	Use multiple performance and accuracy metrics, select and use most appropriate for specific type of data analytics application.	Evaluate and recommend the most appropriate metrics, propose new for new applications.
Define data elements necessary to develop specified data analytics.	Develop specialized analytics to enable decision-making.	Design specialized analytics to improve decision-making.
Choose and execute standard visualization.	Build visualizations for complex and variable data.	Create and optimize visualizations to influence executive decisions.

6 HIGHLY INNOVATIVE TRENDS TOWARDS DIGITAL AND HORIZON EUROPE

Looking into the future, new data-driven technologies (beyond AI which deserves a complete analysis in the 2020 edition of this paper) will be affecting the new work-programs of the EC in the next Framework Program. In the Manufacturing domain, The German Plattform Industrie 4.0 is promoting RAMI as a Reference Architecture and its Assets Administration Shell as a standard (IEC TC 65 WG24) way to model industrial assets along their lifetime (§6.1). In modern collaborative business models, companies need to exchange private data in the value chain, but need to keep absolute control on the access and usage of their data assets, by implementing Data Sovereignty contracts and relevant Digital Enforcement means (§6.2). Lastly, Digital Twins are bridging the gap between real and digital / virtual worlds, enabling advanced cyber-physical applications to be developed and experimented with (§6.3).

6.1 Plattform Industrie 4.0 and RAMI Assets Administration Shell

The German initiative Plattform 4.0 was created to establish standards and guidelines for helping manufacturers adopt key-enabling technologies for Industry 4.0.

RAMI4.0 is a reference architecture model for Industry 4.0. It allows manufacturers to conceive their production systems by deconstructing them through abstraction levels (from top to bottom: business, functional, information, communication, OT/IT integration, and physical asset), granularity levels (from top to bottom: connected world, enterprise, work centers, stations, control devices and field devices) and life-cycles (engineering lifecycle vs. production lifecycle). It also helps to distinguish the viewpoints of product Vs. process. AAS[54] provides a standard means for semantic interoperability, therefore fulfilling the promise of equipment in SMI "to be able to talk to each other". It describes a mechanism for the auto description of the industrial assets; each manufacturing equipment will thus have an "information wrapper" describing its services and the data inputs and outputs it requires to operate. AAS contains the notion of "submodels", which are semantic vocabulary according to industry domains.

RAMI 4.0 and AAS[55] respond to the needs, however challenges for adoption remain:

- **Completion of the AAS standardization efforts:** the DIN SPEC 91345:2016-04, which is the official specification for RAMI 4.0, contains a chapter about AAS and the official data structure (header, body and properties) that an AAS should contain. The later publication from ZVEI (German Electrical and Electronic Manufacturers' Association) goes more into detail about this basic structure. However, this is a meta-structure which does not suffice to deploy AAS in industrial contexts. Concrete properties according to industrial domains should be defined by the submodels in order to deploy AAS. AAS standard submodels do not exist yet. There are propositions/examples on standard submodels, but no standardization body has published them as standards yet. There are different submodel standards that are required:
    - Submodels for functional properties: Submodels that describe what the industrial equipment does, under which conditions, parameters and guaranteed results (e.g. tolerances). This requires a consensus effort in the different industrial communities (e.g. metal manufacturing, plastics, continuous processes).
    - Submodels for non-functional properties:
        - Submodels describing QoS such as safety, security, energy consumption

---

[54] https://www.zvei.org/en/press-media/publications/details-of-the-asset-administration-shell/

[55] https://www.zvei.org/fileadmin/user_upload/Presse_und_Medien/Publikationen/2017/April/Asset_Administration_Shell/ZVEI_WP_Verwaltungschale_Englisch_Download_03.04.17.pdf

        - Submodels containing the proper information to establish communication links through standard protocols. Consensus is needed to choose which protocols will be used where (e.g. standalone OPC-UA for vertical communication, TSN-OPC-UA or M2M for horizontal …)

    Completion of these standardization efforts should allow providers of industrial equipment (e.g. Festo) to provide off-the-shelf usable AAS for their equipment (they might be waiting for the publication of official standard submodels).

- **Tools, Didactic and methodological support:** Despite the widespread dissemination and the quality material that Plattform 4.0 has produced, support for adoption according to industry types remains missing:
    - Tooled methodologies are needed, that allow the definition of a production system which conforms to RAMI 4.0.
    - A showcase of RAMI 4.0 reference implementations per industrial domain.
    - Didactic factories have flourished as a means to raise the level of skills, both digital and operational, to upskill the force to adopt Industry 4.0 technologies. Examples of how RAMI 4.0 uses didactic factories to explain the reference model and how to implement it remain to be seen.
    - Concrete expected benefits for the manufacturing companies, whether increased production yield or lowering of overall production costs.
    - AAS already has a few standard reference implementations (e.g. OpenAAS developed at RWTH, Eclipse BaSyx hosted at the Eclipse Foundation and contributed mainly by DFKI), but these are missing widespread adoption and have not gone beyond the stage of proof-of-concept in factories.
        - Large-scale pilots are needed to really show the benefits of increased semantic interoperability thanks to AAS:
        - Guidelines for implementation,
        - showcase of experiments;
- **Secure, trusted data value chains**: As AAS allows production equipment to be discoverable, new data flows will appear, enlarging the attack surface of production systems. Cybersecurity will be a growing concern, as management of data privacy and confidentiality in data value chains.
    - Links with IDS (see section on IDS): Data spaces to facilitate enterprise data sharing across the value chain. A common "AAS Industrial Data Space" would enable digital interoperability across actors of the value chain while safeguarding data sharing concerns.
    - Blockchain technologies to establish responsibility and accountability of transactions in data exchange between production equipment:
        - Transactions between AAS-compliant equipment;
        - Formal verification of smart contracts to discard malfunctioning data flow chains;
        - Formal verification of AAS and data exchange implementations to prevent zero-day attacks.
    - Privacy concerns: AAS Connected Products might release personal data; manufacturers will need to be guided for compliance to GDPR regulations.

6.2 Data Sharing Spaces and IDSA Reference Architecture for SMI

In light of emerging AI-driven opportunities, the European Commission has been calling for a coordinated and European-wide action[56] to establish a "common European data space" that can enable

---

[56] Member States and Commission to work together to boost artificial intelligence "made in Europe", EC Press Release, Brussels, 7 December 2018, http://europa.eu/rapid/press-release_IP-18-6689_en.htm

seamless data sharing across borders and enable access to large, secure and robust datasets. The BDVA has more recently published its position on the topic, and identified key actions that need to be taken by European policy makers in coordination with industry, research and academic entities to realize the successful development, implementation and adoption of such a pan-European data sharing space[57]. In contrast to existing implementations which serve the needs of a few entities or are confined to just one industry, this space could allow different vertical, cross-sectorial, personal and industrial data spaces to interoperate, offering broader services and experimentation opportunities to all stakeholders, while adhering to European values. Therefore, prominent data sharing platforms being developed in the SMI sector should consider aligning to standards and best practices (interoperability, security, quality and privacy) in order to facilitate the possibility of integrating with external data sharing ecosystems for a mutual benefit.

A major observation that is especially relevant for SMI is that data has increasingly been considered as an asset in its own right, rather than as a means to an end whose primary value lies in the completion or delivery of a process output. With this in mind, metadata needs to adequately represent and retain context for both backward and forward exploration of industrial processes. Thus, a major BDVA recommendation is for data life-cycle management to natively incorporate data sharing events and transaction capabilities at their core. Other known relevant challenges are the difficulties faced when attempting to establish European Industrial Data Platforms in the global market, the lack of data valuation standards in marketplaces, and factors that result in a lack of data sharing trust and motivation. These include legal questions around the concepts of data producers and data ownership, as well as technical challenges such as unsuitable data verification processes, and unreliable data quality.

The above mentioned position paper includes a domain-independent survey of established methods and emerging solutions for improved data exchange along entire value chains (Annex 2 of the paper). Prominent B2B data sharing initiatives include those put forward by the Industrial Internet Consortium (IIC Layered Databus), the IOTA Foundation (IOTA Data Marketplace) and the International Data Spaces (IDS) Association. In particular, the IDS Reference Architecture Model continues to gain momentum as an emerging de facto standard with the potential to rival or set the bar for other international data sharing space solutions, such as those by the Edgecross consortium[58] and the Industrial Valuechain Intiative[59] (both in Japan), and the MadeInCina2025 strategy[60] for the Chinese manufacturing industry. The BDVA recognizes the IDS reference architecture as a prominent solution that could evolve into a fully-fledged data sharing space. Its origins in Industry 4.0 make it particularly suitable for SMI applications.

A major driver for Industry 4.0 is the broad availability of sensors, communication networks and data processing abilities on different layers. Different Reference Architectures (e.g. RAMI 4.0, IIRA) address this, but do not address the requirements for intercompany data exchange and data sharing. SMI requires comprehensive data exchange and data sharing approaches for a Smart Supply Chain and for the Smart Product Lifecycle founded on an architecture ensuring trustworthiness and data sovereignty along the Supply Chain and the Value Chain.

The IDS project addressed this need and provided a Reference Architecture Model[61] (IDS-RAM) and first implementations for an industrial Data Space. The International Data Spaces Association (IDSA) maintains the Reference Architecture Model through the joint efforts of 100 members from various domains and countries. In parallel the IDSA supports the market adoption of the IDS-RAM for products and services, as well as in enabling an IDS ecosystem.

The IDS approach aims at meeting the following strategic requirements:

---

[57] Towards a European Data Sharing Space: Enabling data exchange and unlocking AI potential, BDVA Position Paper, April 2019.

[58] https://www.edgecross.org/en/

[59] https://iv-i.org/wp/en/about-us/whatsivi/

[60] http://english.gov.cn/2016special/madeinchina2025/

[61] https://www.internationaldataspaces.org/wp-content/uploads/2019/03/IDS-Reference-Architecture-Model-3.0.pdf

- **Data usage control**: In line with the central aspect of ensuring data sovereignty, a data owner in the IDS may attach usage policies to its data before it is transmitted to a data consumer. The data consumer may use this data only if it fully agrees to and ensures that usage policy.
- **Decentralized approach**: The architecture of the IDS does not require central data storage capabilities. Instead, it follows a decentralized approach, which means that data physically remains with the respective data owner / data provider until it is transmitted to a trusted party. Thus, the IDS is not a cloud platform, but an architectural approach to connect various, different platforms (both operational and emerging ones). Nevertheless, participants in the IDS may agree on trusted entities offering central data storage, if deemed necessary.
- **Multiple implementations**: The IDS Connector, being a central component of the architecture, is implemented in different versions, which address the specific needs of domains, usage scenarios and different security profiles. Open Source implementations support the market adoption, but the development and provisioning of IDS Core Components is addressed by the market of software and hardware vendors.
- **Standardized interfaces**: Both the architecture of the IDS Core Components and the communication API are subject to standardization.
- **Certification**[62]: The IDS manifests as a distributed network of Connectors, representing multiple data endpoints. Each implementation of the Connector, as well as each organization seeking admission to the IDS, has to undergo a certification process, ensuring trust and reliability across the entire business ecosystem.
- **Data economy**: The IDS Connector allows the creation of novel, data-driven services making use of data apps (i.e., software components providing dedicated data-related service functionality). The participants in the IDS can request these data apps from a public app store, a private App Store or from another Connector for Remote Software Execution.
- **Secure data supply chains**: The IDS aims at enabling secure data supply chains (i.e., networks consisting of data providers and data consumers), ranging from the source the data originates from (e.g., a sensor on an IoT device) to the actual point of use (e.g., an industrial smart service for predictive maintenance).

In April 2019, Version 3.0 of the IDS-RAM was released and the first assessments for the IDS Certification were conducted. The availability of different implementations of the IDS Core Components addressing different domains and usage scenarios is supported by the consortium projects BOOST 4.0[63], Musketeer[64], AMable[65], MIDH[66] and MARKET 4.0[67]. BOOST 4.0 aims to establish a European Industrial Data Space for Big Data based on the IDS-RAM, including computation on the edge, provisioning of Big Data Analytics Apps and visualization. Musketeer focuses on augmenting the value of shared knowledge by making use of machine learning in federated privacy-preserving scenarios. The specific needs of the Additive Manufacturing Ecosystem for SMEs is addressed by the AMable project. Digital Representations of Additive Manufacturing includes a high amount of intellectual properties and have to be protected accordingly. A Clearing House based on Distributed Ledger Technology ensures traceability[68]. The exchange and use of CPS (Cyber-physical system) data along the Smart Supply Chain and the Value Chain of Smart products is in the focus of the MIDIH project. The MIDIH Reference Architecture provides Open Source Building Blocks based on Apache or FIWARE. MARKET 4.0 targets

---

[62] https://www.internationaldataspaces.org/publications/whitepaper-certification/

[63] https://boost40.eu/

[64] https://musketeer.eu/

[65] https://www.amable.eu/

[66] http://midih.eu/

[67] http://market40.eu/

[68] https://www.internationaldataspaces.org/wp-content/uploads/2019/03/Blockchain-Technology-in-IDS_v1_final_Web.pdf

the rise of Data Marketplaces for Industrial Data by providing a Data Marketplace Reference Architecture based on Data Sovereignty and domain specific implementations of Marketplaces based on the IDS-RAM.

6.3 Data-driven Digital Twins

The recent advances in digital technologies like IoT, Big Data, AI and Digital Twins have brought new opportunities for the manufacturing industry. Global awareness of these technologies together with the expectations of a high return on investment triggered many endeavours. However, the fast-growing number of implementations also faced ambiguity, disconnection and lack of compatibility in different application domains.

Reference Architectures (RA) are used to guide the development of the system for system architects and developers. They describe the structure of the system with its elements as well as the interactions between each element and the environment. They are used as general guidelines which are abstracted from various implementations and use cases. RAs also provide common and consistent definitions, vocabulary and taxonomy of the system that facilitates communication and ensures common understanding within all stakeholders.

The RAs applicable to the manufacturing domain offer their solutions under viewpoints such as business, functional, information, communication, integration, asset, system, security. Among them, the information perspective is the aspect that addresses data access, consistent integration of data, and data integrity. The functional aspect enables the formal description of functions and creates a platform for processing data. Besides widely known and adopted IoT-based RAs such as IIRA and RAMI 4.0, there are other architectures for Big Data such as International Data Space Reference Architecture (IDS RA) and National Institute of Standards and Technology (NIST) Big Data Interoperability Framework Reference Architecture (NBDRA) being, or already, developed. In addition, there are also others ongoing initiatives for developing data-driven RAs for Cloud Computing, AI and Digital Twins.

In line with the European research agenda, there are EU Initiatives offering reference models and architectures stemming from research projects. Among them, the FIWARE digital platform provides its “Powered by FIWARE” reference model and implementation as an open-source data-driven platform. AUTOWARE RA aligns cognitive manufacturing technical enablers, such as robotic systems, smart machines, “cloudified” control, secure cloud-based planning systems, and application platforms to provide cognitive automation systems. FAR-EDGE is another EU project that provides a reference implementation and a blueprint solution, as FAR-EDGE RA, for industrial automation based on edge computing. The IoT ARM, architectural reference model was developed by the project partners of the European FP7 research project IoT-A in 2013 to promote high-level interoperability and sharing best practices in functionality and information usage.

A **digital twin**, like a virtual prototype, is a dynamic digital representation of a physical system. However, unlike a virtual prototype, a digital twin is a virtual instance of a physical system (twin) that is continuously updated with the latter’s performance, maintenance, and health status data throughout the physical system’s life cycle. This leads to high precision models representing physical systems not only through their mechanical of geometric values but also through their behaviour. Using the latest advances in Machine Learning and Artificial Intelligence, this concept opens a new dimension of applications for monitoring, simulation and optimization, developing new products and business opportunities.

The characteristics of a Digital Twin can be summarised as follow:

- Real-time reflection: both physical and virtual parts exist in the Digital Twin, the virtual part can keep ultrahigh synchronisation and fidelity to reflect the physical part.
- Interaction and convergence: it happens in both the physical part, virtual part and between physical and virtual parts. In addition, the interaction and convergence between real-time data and historical data can make the Digital Twin data more comprehensible and useful.
- Self-evolution: the Digital Twin can collect and update data in real-time, the virtual part can continually self-improve by comparing it with the physical part in parallel.

The field of applications for digital twins in smart manufacturing is huge, for example, for product design and development, manufacturing assets and process monitoring for shop floor/ factory, products usage monitoring during the product life cycle, etc.

TYPES OF DIGITAL TWINS

A digital twin implements models of the expected behaviour of the cyber-physical process. These models may be based on scientific and engineering knowledge (first principles, physical or white box models) or statistical analysis (data-driven, data science or black-box models). In reality, all models apart from a fully data-driven model are hybrid, since all physical models use uncertain parameters that are derived from the empirical analysis of data. A spectrum of models, from a fully deterministic model to a purely data-driven model can, therefore, be considered. In practice, both models and measurements are uncertain.

Model-based, Data-driven and Hybrid Digital Twins

This means that a digital twin can be classified as being model-based, data-driven or hybrid. A model-based digital twin uses simulation-based methods to describe the behaviour of the industrial systems. During engineering, it is not possible to easily observe the behaviour of the system. However, during the production process, data becomes available that allows the creation of a data-driven digital twin.

In reality, most commonly the model needed for a digital twin is a hybrid. It is often impossible to simulate the full physical behaviour of a system in real-time. Reduced models must be used. Similarly, models must be tuned in real-time so that they continue to reflect the state of the system. Data-driven models must be constrained within their area of applicability. This is best done by considering the physical and logical constraints on the system. The ultimate goal of the digital twin is to benefit from all available information, not to learn from faulty situations, but to generate synthetic data (using a model approach) to simulate faults that may occur in the future. This produces a virtuous situation – the best of both worlds, where data science methods enhance physical models and vice versa.

Digital Twins for Continuous or for Discrete Processes

It should be noted that Digital Twins have clear differences depending on the type of processes being "virtualized". A model of a continuous (or batch) processes is formulated as a system of differential and algebraic equations. Thus, it is common practice to simulate entire oil platforms and chemical plants in real-time using large models of this type. Operator training using these models is industry best practice.

Other types of manufacturing and logistics systems are best modelled using discrete-event models. These are built on difference equations and discrete changes in state for the components in the model. In this case events and measurements trigger transitions in the state of the system. Models can be either deterministic (we can specify the conditions for every transition) or stochastic (we cannot do this, and must rely on observed data).

DIGITAL TWINS, MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

Data scientists and engineers have built computer models of complex machines and manufacturing processes since the dawn of computing. Machine Learning can be applied to a data stream to uncover or discover patterns in the data. Machine Learning can be applied to raw system measurements, results from a simulator or the differences between measured and observed values. This allows the detection of faults in equipment and sensors. Machine Learning can also automate complex analytical tasks by identifying optimal operation methods.

DIGITAL TWINS AND (INDUSTRIAL) IOT

The power of Digital Twins comes from “high fidelity living models”, and these features depend on a high digitization level of the floor plant: high density of sensors and monitored variables, connectivity with machines and robots, efficient networks for high flows of streaming data, contextualized and semantic data for an efficient machine processing, and a high level of integration.

Here, the use of new technologies such as the Industrial Internet of Things (IIoT) and the Cyber-Physical Systems (CPS), which break this rigidity through the use of standard protocols, IP-based technologies, known semantics and ontologies, and a set of IP-based new generation networks, comes into play. However, these technologies are still young, immature and at an early phase of adoption and will require considerable customization work to create bespoke digital twins for their unique operational needs.

Augmented Reality (AR) is one of the key technologies in Industrial 4.0 and in particular it can be used to implement the digital twin for industrial processes. Users can take advantage of that virtual model to perform efficient decision-making and higher-level machine control. Therefore, this not only improves the efficiency and effectiveness of manufacturing but also raises the connection between the virtual part and the physical part of Digital Twin to another level.

CALIBRATION AND VALIDATION OF A DIGITAL TWIN

Effective calibration and validation of models is essential to build successful digital twins. One of the main reasons is the dynamicity of the environment the physical assets are living in. This is a fertile area for taking advantage of recent developments in statistical optimization and machine learning. For design, production, and timely marketing it is, therefore, necessary to understand all individual components and the overall system itself at all stages of the development process in theory and practice. Combining engineering processes with appropriate simulation techniques allows the comparison and evaluation of different system designs and to reduce the number of physical prototypes.

CHALLENGES IN DIGITAL TWINS

In this section, eight challenges that need to be addressed to allow digital twins to release their promised potential will be detailed as follows:

1. **Business Models, Security and Confidentiality.** All participants in the supply chain can benefit from digital twins. An operator has a strong incentive to have a comprehensive twin for monitoring and optimization. Vendors can benefit from using operational data in their product twins. This kind of data sharing and collaboration approach for simulation and analytics is of undeniable value for all the involved actors. However, mechanisms must be found for fair rewards, allocating responsibility, ensuring secure access and protecting IPR.
2. **Work practices.** The digital twin can change work practices, but only if it offers the users tangible and measurable benefits. The user must be convinced that the system gives benefits, that it's safe and usable.
3. **Scope.** "Fit for purpose modelling", and the principle of parsimony (i.e. just enough for functionality, but no more) is required. Models with a different granularity that allow a user to zoom on the digital twin as needed should be supported.
4. **Usability.** A comprehensive digital twin has multiple concerns and can produce massive amounts of data. A specialized user must be able to find the needed information quickly, easily and without distraction.
5. **Integration.** A digital twin will consist of many data sources and simulation models. An operating company will have to cope with several vendor platforms, either proprietary or open, in addition to their own, internal legacy systems.
6. **Maintenance.** Life Cycle Management of a Digital Twin is one the main challenges, due to the complexity of software and hardware systems. It combines life-cycle information, measurements of the asset state and simulations. All these must be maintained so that they reflect the as-built state of a system. This includes new processing, like cognition, leading to self-aware twins that are able to keep them efficient and valid.
7. **Computational overload, edge and cloud.** A comprehensive Digital Twin will require extensive computational resources probably distributed over a hybrid cloud that combines private clouds with vendor platforms and HPC resources. This design will need to carefully define what should be done at the edge of the system and what is done in the cloud.
8. **Uncertainty, Validation and Data Science.** Finally, a digital twin is only as good as the data and models used in the system. Data must be cleaned and reconciled. Models (event "first principles") must be validated and tuned to ensure that they follow the state of the facility.

A RESEARCH AGENDA FOR DIGITAL TWINS

In order to have Digital Twins sustainable and maintainable, it is necessary to combine the sub-disciplines of knowledge representation, natural language technologies, formal methods, scalable computing and data science. This knowledge of technologies must be informed by the deep domain knowledge that is embedded in the digital twins' simulation models and is owned by the facility's engineers – chemical, petroleum, mechanical, electrical and control – and managers. Finally, there remains the challenge of uncertainty, validation and data reliability. A digital twin is built on models, usually many models. Some are based on physical principles: structural, geometrical and process simulations. Others are purely empirical, based on machine learning. These models must be validated against observed facility behaviour and aligned so that they mirror observed normal behaviour. Aligning models to observed data is difficult and remains an art. A maintainable digital twin will contain structured tools that allow validation and tuning of all the models in the system. We believe that hybrid analytics – the combination of data science with physical and engineering simulations – is a valuable and fruitful area of research. Machine learning can benefit from being constrained by the laws of physics, while the laws of physics contain parameters that are uncertain or expensive to measure. Good statistical practice is needed in the engineering communities and engineering knowledge is needed among data scientists.

7 RECOMMENDATIONS AND FUTURE OUTLOOK

7.1 Recommendations for materialising SMI Grand Scenarios

Based on the feedback from challenges analysis in the three grand scenarios by the BDVA community, we present in the following a short list of recommendations to drive the development and deployment of BD Technologies in Smart Manufacturing Industry.

R1. **POLICY**. The European Research & Innovation approach to AI and "Big Data Technologies for Smart Manufacturing Industry" needs to be driven by compliance with EU legislation, principles and values, strongly supporting a human-centric approach. It shall be grounded in the rule of law and in the protection of fundamental rights, such as human dignity, non-discrimination, freedom of expression and privacy protection. In line with the EC's White Paper on Artificial Intelligence - A European approach to excellence and trust[69], EU SMI will play a key role in achieving the Sustainable Development Goals and in supporting the societal wellbeing, aligning its further research and deployment actions with the EU objectives to achieve an "ecosystem of excellence" coupled with an "ecosystem of trust". The requirements set by the Ethics Guidelines for Trustworthy AI, including the input obtained during their piloting phase, need to be followed, paying attention also to the transparency, traceability and human oversight (not specifically covered under current legislative regimes in several sectors). At the same time, in view of supporting the EU efforts towards the upward regulatory convergence, the OECD's ethical principles for AI and the insights coming from other multilateral fora need to be taken into account.

R2. **TECHNOLOGY SMART FACTORY**. Expand the current portfolio of ontologies and taxonomies (mostly developed for materials) to specific areas of manufacturing not covered yet and make it open for the research community. Develop the interoperability of the current set of Data Platforms from R&D projects to reduce fragmentation and avoid reluctance of usage from end users, especially SMEs. Continue promotion of European AI-on-demand platforms to favour its maturity and development with AI assets (datasets, algorithms and applications) populated from FP projects related with AI for manufacturing.

R3. **TECHNOLOGY SMART PRODUCT**. Thanks to Digital Twins and Digital Personae enablers, the next Framework Program should develop and deploy a seamlessly interoperable Data Sharing Space for Things and Humans along the lifecycle of Products, Data and Services. Industrial Data Platforms will be able to semantically interoperate highly heterogeneous data sources belonging to different administration and security domains along the lifecycle of complex products.

R4. **TECHNOLOGY SMART SUPPLY CHAIN**. The supply-chain relevant data is characterized by significant sensitivity, so that its collection, storage and further processing beyond company boundaries requires considering security-relevant standards. In order for companies in supply chains to continue to secure and expand their value creation, it is necessary to ensure data sovereignty. This requires the development of new technical concepts for data sovereignty and security by means of data flow control, trusted data networks, data provenance and the administration and distribution of usage restrictions. The proliferation of these technologies will significantly improve the smart supply chains, as the companies will be able to share the data in a way that enables monitoring and inspires trust. In this context, the IDS offers an excellent reference architecture and a set of services that should ensure the necessary data sovereignty.

R5. **BUSINESS MODELS**. Promote the development for the manufacturing industry of a Data Economy and the development and deployment of a fully European Industrial Data Platform to support new data models leveraging the benefits of servitisation. Extracting value from industry-

---

[69] COM (2020) 65 final (19th February 2020).

generated data means to create new business opportunities which are not just related to the manufacturing industry in itself, but it is also able to mobilise resources and ecosystems of advanced solution providers, which can increase the competitiveness of our Manufacturing Industry in the global competition arena.

R6. **SKILLS**. Promote the development of an EU competence framework and an EU observatory for skills needs assessment and forward looking in relation to big data and artificial intelligence for manufacturing. Create EU international degrees and masters, modernize technical & vocational education and training, promote training-on-the-job programmes on big data and AI in manufacturing. To this end, collaborative schemes based on industry-research-education-government alliances must be fostered. Exploiting novel methods and delivering mechanisms within new “cyber-physical” learning spaces to develop big data and AI-related competences and bring them closer to industrial practice, should be promoted. Documenting skills and qualifications in more transparent and portable ways – e.g. digitally-signed credentials - to facilitate the match between data science and AI skills supply and demand in manufacturing, and beyond it, should also be encouraged.

7.2 Future Outlook and 2020 Plan

The 2020 edition of the present document will represent the bridge between H2020 research and innovation and the next Financial Framework Program in 21-27. The 2020 discussions in the BDVA SMI subgroup will be focussing, indeed, on the following four main pillars:

- **Artificial Intelligence**, where the AI paradigm will be mostly directed towards: the integration of Manufacturing in the AI on demand ICT-26, ICT-48 and ICT-49 community; the peculiarity and commonality of AI for manufacturing applications also in the Process Industry domain (projects CAPRI and COGNITWIN); the new role and competencies required by humans to interact with autonomous systems as well as the Collaborative Intelligence interaction paradigm; the development of AI for manufacturing testing and experimental facilities.
- **Data Spaces and Platforms**, where the BDVA Position Paper about “Towards a European Data Sharing Space” will be discussed among the “BDT for Manufacturing” projects, with special focus on the new Industrial Data Platforms ICT-13 and their industrial cases. The IDS Reference Architectural Model will be also followed, commented from the SMI perspective and its reference implementations and use cases analysed against the chosen architectural and technical solutions in the Industrial Data Platforms projects.
- **Secure computation infrastructures for BDT,** where fundamental links will be established between BDVA and the community of ECSEL and the Industry4.E CSA in the domain of Smart Systems; the edge-cloud-HPC community for a continuum computational space and in particular the ECSO community for trusted and secure data infrastructures. The evolution of the trilateral (Germany, France, Finland) GAIA-X initiative for a European Digital Sovereignty will be also followed with particular attention.
- **Digital Innovation Hubs and SMEs Digital Transformation,** where BDT for Manufacturing challenges will be weighed in the context of SMEs and proper customisations, instantiations delivered. In particular, the 2020 paper will listen to the voice of the DIH-SAE community (DT-ICT-01) and especially its projects about Cyber Physical Systems, of the DIH-I4MS Phase IV (DT-ICT-03) which will be launched towards the end of the year and the currently running DT-ICT-02 about Agile Production in Robotics and the two Innovation Actions TRINITY and DIH2. Constructive relations will be also built with the VANGUARD Initiative and the Efficient and Sustainable Manufacturing pilot as well as with the evolution of the AI DIH Network initiative led by PwC.

ANNEX

Editorial Team and Contributors

Editors

- Davide Dalle Carbonare (Engnieering)
- Sergio Gusmeroli (POLIMI)
- Anibal Reñones (CARTIF)
- Oscar Lazaro (INNOVALIA)
- Ljiljana Stojanovic (Fraunhofer Institute IOSB)

Contributors

Gabriella Cattaneo (IDC), Georgios Karagiannis (HUAWEI), Orlando Avila-García (Atos Spain), Ernoe Kovacs (NECLAB), Nenad Stojanovic (NISSATECH), Natalie Bertels (KU-LEUVEN), Jon Kepa Gerrikagoitia (IDEKO), Marta Pinzone (POLIMI), Almudena Sánchez González (GMV), Ernestina Menasalvas (UPM), Juan Cadavid (CEA), Daniel Alonso Roman (ITI), Simon Scerri (Fraunhofer Institute IAIS), Nuria Gomez (GMV), Perin Unal (TEKNOPAR), Chris Decubber (European Factories of the Future Research Association - EFFRA), Marina Cugurra (GFT Italia), Nicholas S. R. Fair (IT Innovation) Sebastian Steinbuss (IDSA).

In addition, special thanks to all BDVA members of the Smart Manufacturing Industry group. The content of this paper has been produced also thanks to the active collaboration in various events (BDVA AG meetings, BDV PPP Summits, EBDVF events).

Note

This document should be referenced as follows: *“Gusmeroli S., Dalle Carbonare D. (eds) (2020) Big Data challenges in Smart Manufacturing Industry (v. 2020). Brussels. BDVA”*

BDV BIG DATA VALUE ASSOCIATION

www.bdva.eu