Ref. Ares(2021)1663942 - 05/03/2021

D 2.2
Pathways cross-fertilisation with Digital Technologies- First Iteration v.1.0

Action acronym	ConnectedFactories 2
Action Full Title	Global-leading smart manufacturing through digital platforms, cross-cutting features and skilled workforce
Grant Agreement Number	873086
Instrument	CSA: Coordination and Support Action
Deliverable Number	D2.2
Deliverable Title	Pathways cross-fertilisation with Digital Technologies - First Iteration
Lead Beneficiary	CEA
Work package	2
Work package leader	FPM
Dissemination level[1]	Public
Type[2]	R
Due date according to DoA	M15
Actual submission date	5.3.2021
Main editors and contributors:	FPM, LMS, INNOVALIA

VERSION MANAGEMENT

[1] PU: Public, CO: Confidential, only for members of the consortium (including the Commission Services)
[2] RE: Report, OT: Other; ORDP: Open Research Data Pilot

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

	Name	Beneficiary
Author(s):	Ariane Piel, Juan Cadavid	CEA
Contributor(s):	Marta Pinzone, Sergio Gusmeroli, Kosmas Alexopoulos, Jesus Alonso	FPM, LMS, INNOVALIA
Reviewed by:	Kosmas Alexopoulos	LMS

Abbreviations and acronyms

TERMS, ABBREVIATIONS AND ACRONYMS	
AI	Artificial Intelligence
B2B	Business-To-Business
DEP	Digital Europe Program
CF	Connected Factories
DGA	Data Governance Act
DFA	Digital Factory Alliance for Digital Manufacturing Platforms
IDSA	International Data Spaces Association
PPP	Public Private Partnership

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

TABLE OF CONTENT

Executive Summary 5
1 Introduction 7
1.1 The CF1 Pathways and the future of Connected Factories 8
2 Data-driven Digital Technologies trends and roadmaps 12
2.1 The Data Revolution and Digital Manufacturing Platforms 12
2.1.1 The Data Governance Act DGA 14
2.1.2 Roles, Professions and Skills and Data for Manufacturing 15
2.2 The AI Revolution and Digital Manufacturing Platforms 17
2.3 The Digital Factory Alliance for Digital Manufacturing Platforms 20
2.3.1 Excellence in data-driven factory 20
2.3.2 DFA Goals 21
2.3.3 DFA 2021 Plan 22
3 Business aspects towards a European Data Space for Manufacturing 23
3.1 The Market Perspective: from data-powered enterprises to Data4AI perspective 23
3.1.1 The Data-powered Enterprise (CapGemini) 23
3.1.2 The Data for AI Challenge (Sloan MIT Review) 27
3.2 Business Models for Data Spaces in the Private Sector 29
3.2.1 Open Data in Manufacturing 30
3.2.2 Data Marketplaces in Manufacturing 32
3.2.3 Trusted Data Networks in Manufacturing 34
4 Materialising Data Spaces in the Manufacturing Sector 36
4.1 The European Commission perspective 36
4.2 The BDVA Perspective: Towards a common European Data Sharing Space 39
4.3 The International Data Spaces Association perspective 41
4.4 The OPEN DEI perspective: Task Force 1 Design Principles for Data Spaces 44
4.5 A Reference Matrix for Data Spaces in Manufacturing 47
5 The CF2 Data Spaces pathway 48
5.1 Data Spaces levels and Dimensions 48
5.2 The Data Spaces Pathway Survey 49
5.3 The Data Spaces Pathway Interactive Board (MURAL) 52
5.4 Validation and Experimentation of the Data Spaces Pathway 53
5.4.1 BOOST4.0 Experimentation 53
5.4.2 DMP Cluster Experimentation 54
6 Conclusions and Future Outlook 55

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

TABLE OF FIGURES

Figure 1 - Autonomous and Smart Factory Pathway and Key Enablers 9
Figure 2 - Autonomous and Smart Factory Pathway 10
Figure 3 - Collaborative Product-Service Factories Pathway 10
Figure 4 - Hyperconnected Factories Pathway 11
Figure 5 - Reference framework for EU Digital Manufacturing 12
Figure 6 - EU main pillars for Data Strategy 13
Figure 7 - European Data Space 14
Figure 8 - Pillars of the EU Data Governance Act of November 2020 15
Figure 9 - 5 main job profiles for Data Space management 16
Figure 10 - European AI, Data and Robotics Framework and Enablers released in September 2019 17
Figure 11 - Six Technology Enabler interactive workshop 19
Figure 12 – How many enterprises make Data-driven decision, by sector (The data-powered enterprise, August 2020) 24
Figure 13 – CapGemini evolutionary Data pathway for enterprises (The data-powered enterprise, August 2020) 25
Figure 14 - Capgemini four classes of Data exploitation for enterprises (The data-powered enterprise, August 2020) 26
Figure 15 - AI exploitation framework, Sloan MIT Review 28
Figure 16 - SCSN structure for manufacturing data exchange 33
Figure 17 - Boost4.0 FIAT use case of Data Exchange in a close platform 35
Figure 18 - Boost4.0 WHIRPOOL use case of Data Exchange in an open platform 35
Figure 19 - EU study about barriers and challenges for Data Share and Data Re-use, November 2020 37
Figure 20 - European Big Data Value Forum: Model for manufacturing Data Spaces 38
Figure 21 - Manufacturing Data Spaces, DEP workprogramme (2021-2022) 38
Figure 22 - The European Data Sharing Value Wheel 40
Figure 23 - IDSA layered architectural model 42
Figure 24 - Data Space elements from the point of view of Connectivity, Trust and Governance 42
Figure 25 - IDS reference implementation model 43
Figure 26 - High level overview of the GAIA-X architecture 43
Figure 27 - OPEN DEI study on European Data Sharing Space 46
Figure 28 - Reference matrix for Data Space in Manufacturing 47
Figure 29 - Data Space Evolutionary Matrix 49
Figure 30 – MURAL Interactive Workshop Board about DS Pathway 53

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

Executive Summary

Deliverable D2.2 (M15) "**Pathways cross-fertilisation with Digital Technologies - First Iteration**" is the first release of CF2 cross-fertilisation with new trends and roadmaps from the domain of Digital Technologies. In the spirit of WP2 "Pathways to digitalisation of manufacturing" this means to take inspiration from Digital Technologies to develop new pathways for Manufacturing Industry which could measure the readiness of enterprises with respect to the full adoption and implementation of such technologies.

More in particular, T2.2 Pathways cross-fertilisation with Digital Technologies, says. "*Using a commonly adopted metaphor, manufacturing systems and digital technologies can be regarded as the two sides of the same Industry 4.0 coin. This task will materialise or explicate the coin metal between the CF pathways side and the enabling and emerging digital technologies side, by identifying, supporting and implementing osmosis processes between the two domains. More in particular, digital innovations, such as in the AI or Cybersecurity domains, will be progressively and selectively (osmosis) transferred to the manufacturing domain, by enhancing and extending the current pathways by such new technologies. Conversely, new features of the manufacturing domain pathways could indicate new requirements and specifications for the digital technology side, as for instance required by the circular economy principles, humans in the loop factors and non-hierarchical manufacturing ecosystems*".

In this first release of D2.2 deliverable, we concentrate on the so-called "**Data Revolution**", i.e. how Data Economy could affect and influence the business of Manufacturing companies and how the Manufacturing industry is ready to fully adopt the potential benefits of this revolution. In the following release, we will be focusing especially on AI, Cyber-security and HPC/Cloud/Edge in line with the main technologies at the basis of next 2021-2027 Digital Europe Programme.

Data Technologies are at the basis of any Digital Transformation both at technological and at business level. This has been depicted, for instance, by the collaborative product-service pathway in CF1 and the servitisation business models deriving from it.

From a **technological offer-side viewpoint**, Data Technologies are evolving to AI (see the Data4AI challenge in the SRIDA of the new AI Data Robotic DAIRO Partnership) and the main challenge for Manufacturing Industry is to be able to provide AI experts with high quality meaningful datasets (the Garbage-in Garbage-out metaphor). IDSA (International Data Spaces Association) was indeed born as Industrial Data Space Association with a strong focus on Reference Architectural Models for implementing B2B Data Sovereignty. Now the clear focus is on Data Spaces Building Blocks identification (the soft infrastructure made of software, models and governance rules) and the OPEN DEI white paper on "Design Principles for Data Sharing Spaces" (to be published soon) is an attempt to define agreed common rules among the sectors to implement and materialise Data Spaces. Intense discussions are currently in place about the Maturity of such Building Blocks and whether it is timely to open right now in 2021 sector specific Data Spaces in the DEP programme, while many of such Building Blocks are still at an embryonic development stage (see for instance the Usage Control models and tools and the integration of standard interoperability models for data payloads).

The D2.2 is focussing on the **Demand side** of the Data Spaces coin. Several market analyses show on the one side the exceptional potential of the Manufacturing market for Data and AI Technologies, but on the other side a current insufficient maturity of many manufacturing enterprises (especially SMEs) in taking decisions based on data analytics and decision making (the so-called data-powered enterprises). CF2 Task 2.2 is providing an evolutionary model and a set of interaction tools (self-assessment and participative living lab

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

assessment) for SMEs to position themselves in a 5 levels Data Maturity scale and to plan concrete actions to implement a stage-by-stage migration pathway.

The CF2 Data Spaces pathway includes six analysis dimensions:

a. ***Data Management****: Principles and techniques for data lifecycle management including quality, integrity, provenance and liability*
b. ***Data Protection****: Privacy and anonymisation mechanisms to facilitate data protection. This is shown related to data management and processing as there is a strong link here, but it can also be associated with the area of cybersecurity.*
c. ***Data Processing Architectures****: Optimised and scalable architectures for analytics of both data-at-rest and data-in-motion, with low latency delivering real-time analytics*
d. ***Data Analytics****: Data analytics to improve data understanding, deep learning and meaningfulness of data, supporting forecasting, prediction, prescription and autonomy*
e. ***Data Visualisation and User Interaction****: Advanced visualisation approaches for improved user experience*
f. ***Data Sharing along Value Ecosystems*** *and Innovation Support: Data platforms for data sharing include, in particular, Industrial and Personal Data Platforms, but also other data sharing platforms like Research Data Platforms (RDPs) and Urban/City Data Platforms (UDPs).*

These six dimensions have been chosen for the analysis of "Data Space" maturity of a Manufacturing company (SME), along five levels of maturity defined as:

i. **No Data Control**. This is a level where companies do not have CONTROL of DATA along their lifecycle.
ii. **Data Silos.** This is a level where the company understands the relevance of Data lifecycle management as for feeding Enterprise applications such as PLM ERP SCM CRM MES., but in any case DATA is SILOED.
iii. **Data Bridges.** This is a level where for feeding some advanced applications in the domain of planning, optimisation, prediction, the company needs to integrate data from different sources (e.g. ERP-CRM; ERP-MES, PLM-MES).
iv. **Data Interoperability.** This is a level where enterprises define a strategy and a set of processes for managing and governing the Data generated inside and acquired from outside along their complete lifecycle.
v. **Data Valorisation.** This level is when the enterprise fully understands how data can be profitably exploited inside and outside of the value chain. Data Marketplaces and/or Trusted Data Networks are setup and managed in agreement with Data Protection (personal data) and Data Sovereignty (non-personal data) principles.

Now the DS pathway 6*5 contribution space is being filled by contributions from projects and especially ZDM projects are required to provide its view in the domain of Data Spaces for Zero Defect Manufacturing in Autonomous Quality scenarios. ZDM industrial pilots will be soon required to compile an online survey and to be interviewed about their maturity regarding Data Spaces. Such a maturity analysis could also be interpreted as an action to improve Data Culture and Data Value Awareness also in the manufacturing industry, where Data Economy principles are now leading new business and innovation models in other domains such as Mobile Communication, Retail and Finance.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

1 Introduction

Deliverable D2.2 (M15) "Pathways cross-fertilisation with Digital Technologies - First Iteration" is the first release of CF2 cross-fertilisation with new trends and roadmaps from the domain of Digital Technologies. In the spirit of WP2 "Pathways to digitalisation of manufacturing" this means to take inspiration from Digital Technologies to develop new pathways for Manufacturing Industry which could measure the readiness of enterprises with respect to the full adoption and implementation of such technologies.

More in particular, T2.2 Pathways cross-fertilisation with Digital Technologies, says. "*Using a commonly adopted metaphor, manufacturing systems and digital technologies can be regarded as the two sides of the same Industry 4.0 coin. This task will materialise or explicate the coin metal between the CF pathways side and the enabling and emerging digital technologies side, by identifying, supporting and implementing osmosis processes between the two domains. More in particular, digital innovations, such as in the AI or Cybersecurity domains, will be progressively and selectively (osmosis) transferred to the manufacturing domain, by enhancing and extending the current pathways by such new technologies. Conversely, new features of the manufacturing domain pathways could indicate new requirements and specifications for the digital technology side, as for instance required by the circular economy principles, humans in the loop factors and non-hierarchical manufacturing ecosystems*".

In this first release of D2.2 deliverable, we concentrate on the so-called "**Data Revolution**", i.e. how Data Economy could affect and influence the business of Manufacturing companies and how the Manufacturing industry is ready to fully adopt the potential benefits of this revolution. In the following release, we will be focusing especially on AI, Cyber-security and HPC/Cloud/Edge in line with the main technologies at the basis of next 2021-2027 Digital Europe Programme.

At any rate, Data Technologies are at the basis of any Digital Transformation both at technological and at business level: let's think for instance at the collaborative product-service pathway in CF1 (see §1.1) and the servitisation business models deriving from it.

In this introductory chapter, we will make a synthesis of the whole document and a reminder of the CF1 pathways which as reference North Stars will drive all the CF2 activities (§1.1).

Chapter 2 aims at describing the EU Data Strategy of February 2020 (§2.1) with the lenses of Manufacturing Industry, which is one of the nine critical sectors where the EC sees the most relevant advantages for EU Business and Society. Further paragraphs in Chapter 2 will be moreover devoted to other Digital revolutions such as Artificial Intelligence (§2.2). Finally, the Digital Factory Alliance (§2.3) will be described as a tangible initiative to unify all Digital Technologies for Manufacturing in a one-stop-shop and community.

Chapter 3 will start analysing the impact of the Data Revolution on the Manufacturing Industry and in particular the Digital Manufacturing Platforms. The concept of Data-Powered enterprise will be introduced and discussed (§3.1) as well as the three categories of Business Models for the Private sector identified by the EC as enabled by the implementation of a Data Strategy, namely Open Data, Data Marketplaces and Trusted Data Networks (§3.2).

The overall high-level scenario depicted in Chapter 3 will be then materialised in **Chapter 4**, where the notion of Data Space for Manufacturing will be introduced and discussed. In a series of dedicated workshops (the last one held on November 23rd 2020) several stakeholders expressed their opinions about opportunities and challenges for implementing Data Spaces in Manufacturing. The European Commission (§4.1), the BDVA (§4.2), the International Data Spaces Association (§4.3) and the OPEN DEI CSA (§4.4) expressed their viewpoint and paved the way towards a European way to implement the Data Revolution in Manufacturing.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

While in Chapter 4 the Offer side will be studied and discussed, **Chapter 5** is fully devoted to the Demand side, in particular to understand barriers and challenges in the different processes of the Manufacturing Industry for the full adoption of Data Technologies. The "Data Spaces pathway" developed in CF2 WP2 and described in Chapter 5 is an original methodology to assess the readiness of a company to embrace Data Technologies in their products, processes and people mindset, from the workers to the C-level decision makers. A bidimensional Matrix (§5.1) has been developed by crossing the BDVA technology enablers with five levels of maturity spanning from a total un-awareness of the business potential of Data to a full exploitation of the Data Value. An online survey (§5.2) and an Interactive Whiteboard (MURAL §5.3) have been developed and validated (§5.4) with some industrial cases coming from EU projects (especially from Boost4.0 Big Data Lighthouse project for the Manufacturing Industry).

Further validation of the methods and tools will be conducted on selected DMP cluster projects and reported in the next edition of this deliverable, where also other "digital revolution" aspects will be analysed under the lenses of the Manufacturing critical processes.

1.1 The CF1 Pathways and the future of Connected Factories

ConnectedFactories CSA[3] has been conducted along 2017-2019. In this section, the latest evolution of the three Big Data Technologies adoption pathways will be reported. Pathways to the digitalisation of manufacturing reflect how digitalisation (data driven in particular) and eventually the deployment of digital platforms (industrial data 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.

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**.

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

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

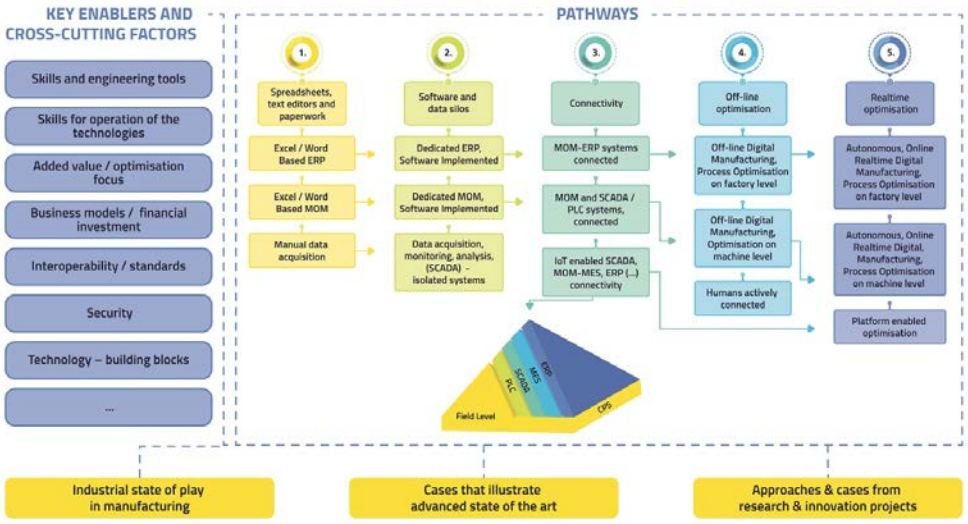


Figure 1 - Autonomous and Smart Factory Pathway and Key Enablers

Moreover, there are many **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[4] 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 viewpoint**. 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 companies should be highlighted**. It should be noted however that large manufacturing is also facing challenges in making their way along the pathways.

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

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

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. With respect to value creation resulting from big data, the **Autonomous & Smart Factories pathway** (see picture below) 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).

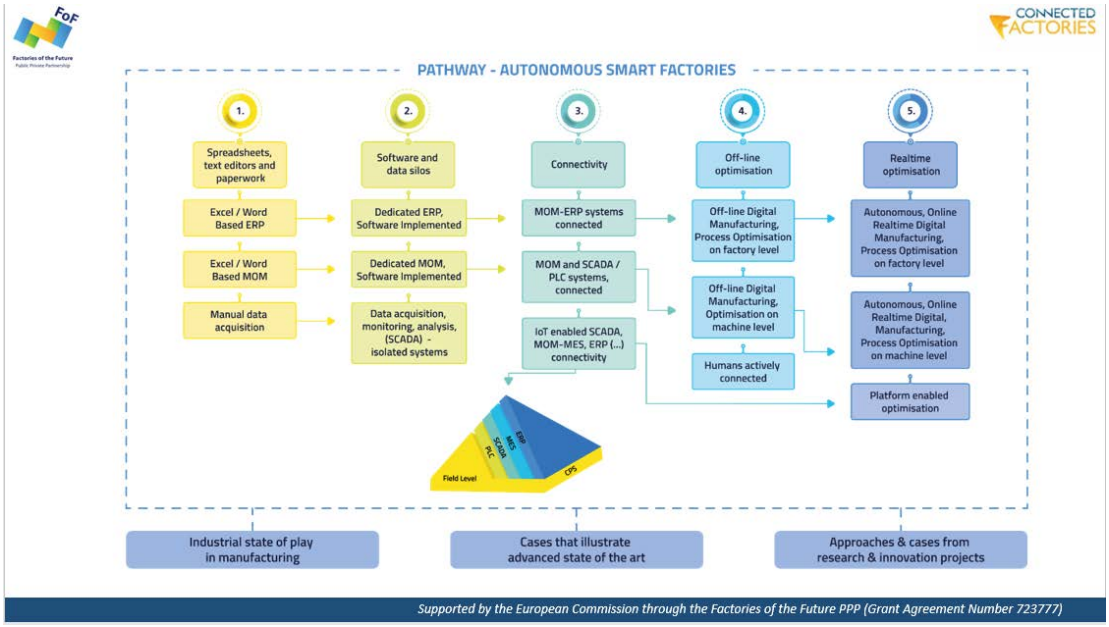


Figure 2 - Autonomous and Smart Factory Pathway

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).

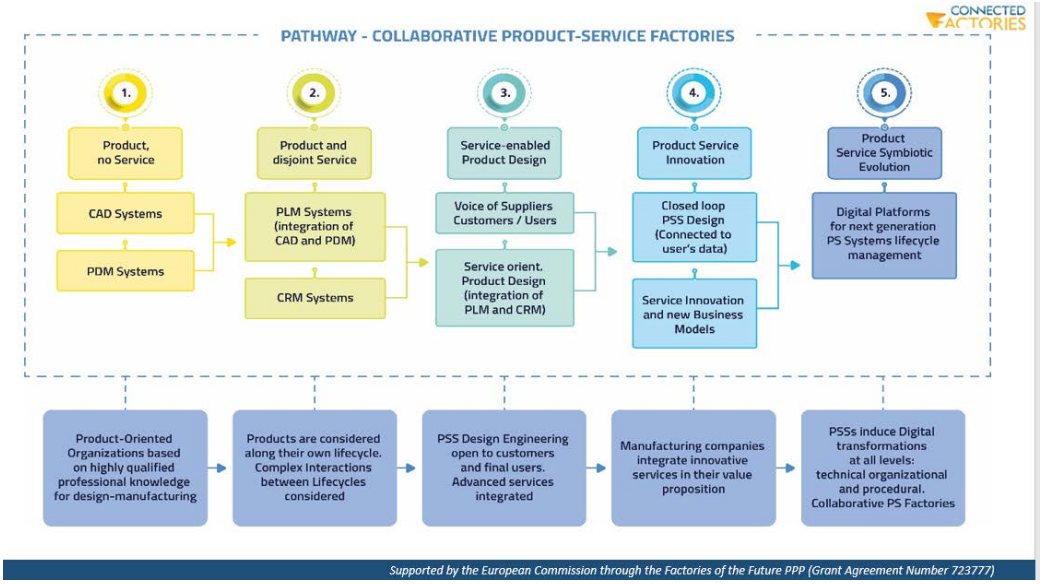


Figure 3 - Collaborative Product-Service Factories Pathway

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

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).

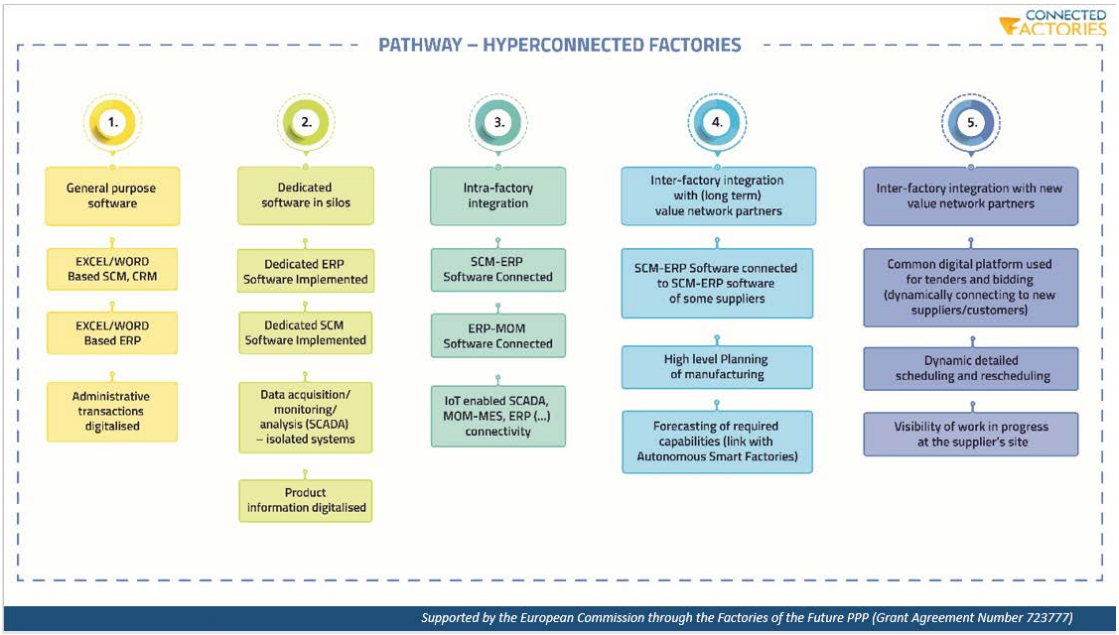


Figure 4 - Hyperconnected Factories Pathway

A Data-oriented interpretation of the three pathways shows unprecedented opportunities and challenges, which will be deeply investigated and analysed in the rest of the present deliverable.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

2 Data-driven Digital Technologies trends and roadmaps

This first release of D2.2 will focus on Data Technologies revolution and the impact on Manufacturing Industry Products, Processes, Organisation and Competencies. In §2.1 the EU Data Strategy will be analysed with its four pillars and its implications to our domain sketched. In the evolutionary pathway from Data to AI, §2.2 will report about Data4AI challenges and how the Data Deep Dive of the "Data, AI, Robotics Partnership" Strategic Research, Innovation and Deployment Agenda (DAIRO SRIDA) has been recently commented and discussed inside the community of Manufacturing experts of the BDVA SMI group, which is closely related to CF2 WP2. Very recently, inside the DMP Cluster and in particular in the context of the QU4LITY and of the Boost4.0 projects, a new initiative has been launched (the Digital Factory Alliance) with the aim to create a community and one-stop-shop for data-driven Factories of the Future, closely linked to the CF1 Smart Autonomous Factory pathway.

2.1 The Data Revolution and Digital Manufacturing Platforms

The need for a common EU Data Space for Manufacturing Industry was recently highlighted by different initiatives and communications. In the field of Important Projects of Common European Interest (IPCEI[5]) at the beginning of 2019 a reference framework for EU common Industrial IOT, Data and AI ecosystem was presented, where the first pillar of the Digital Platform enabling is called: "**Secure and Trusted Data Spaces: data access, sharing and valorisation**". Manufacturing is one of the verticals where such an IPCEI is going to be developed.

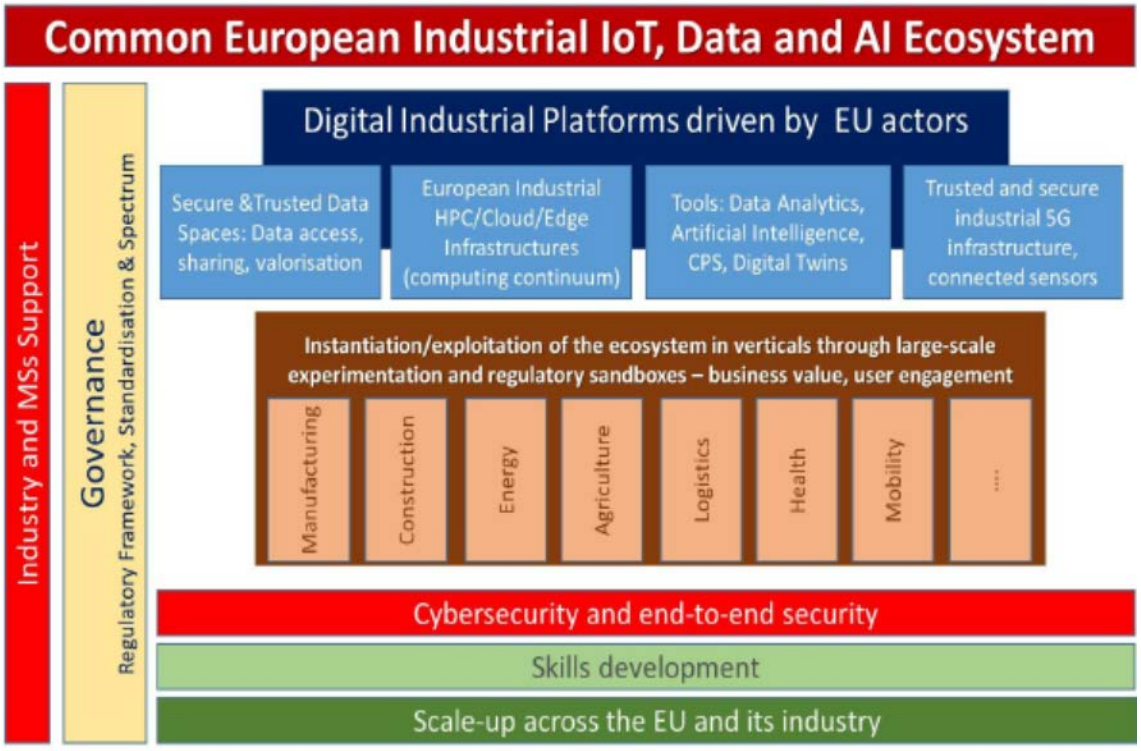


Figure 5 - Reference framework for EU Digital Manufacturing

[5] https://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupMeetingDoc&docid=28145

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

More recently, the **EU Data Strategy**[6] of February 2020 has identified the four main pillars along which such a strategy will be implemented in the next few years: a new cross-sectorial **governance** framework for data sharing, a new technology development framework (**enablers** and building blocks), a new data-oriented education and **skills development** program and a set of **vertical rollouts** in crucial economic and business sectors.

Manufacturing is one of such crucial sectors where Data Spaces are being developed and deployed in several scenarios such as **Smart Autonomous Factory, Hyper-connected Factory and Collaborative Product-Service Systems**.



Figure 6 - EU main pillars for Data Strategy

PILLAR I has been recently implemented by the EU Data Governance Act[7] (§2.1.1) which is now defining the legal and regulatory issues driving the European way to Data Spaces.

PILLAR II represents the technical aspect of a European Data Space, including the integration of FAIR Data Sets (metadata, standard data models, ontologies) with Industrial Data Platforms under a common set of governance and behavioural models (access and usage rights models). This will be described in detail in **Chapter 4** of this deliverable D2.2.

PILLAR III addresses the point of specific professions, roles and competencies needed for the Data Revolution, which is fully addressed by CF2 WP4 and in particular D4.1 (M15 as well as this D2.2). A short summary of CF2 findings in this perspective is here reported in §2.1.2.

Finally **PILLAR IV** is the **main subject of this deliverable** regarding the Manufacturing Industry, according to the picture here below. On top of a technical infrastructure for data spaces (e.g. based on IDS Reference

[6] https://ec.europa.eu/digital-single-market/en/policies/building-european-data-economy
[7] https://ec.europa.eu/digital-single-market/en/news/proposal-regulation-european-data-governance-data-governance-act

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

Architecture and GAIA-X infrastructure), a rich pool of High Value Data Sets (by Volume, Velocity, Variety, Veracity) has to be managed by (personal and industrial) Data Platforms. In the picture nine crucial economic sectors have been identified and "Industrial & Manufacturing" is one of the most relevant sectors when an EU common Data Space is required. A business-oriented analysis of "Data for Manufacturing" impact is reported in Chapter 3.

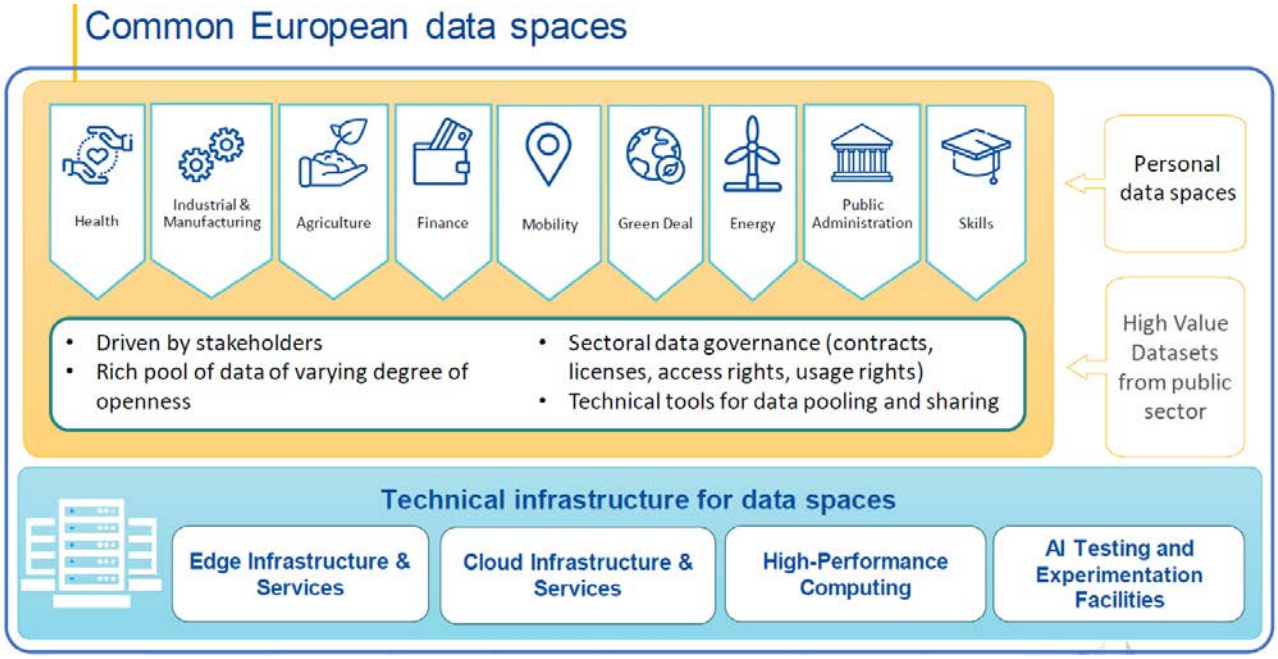


Figure 7 - European Data Space

2.1.1 The Data Governance Act DGA

The **Data Governance Act**[8] of November 25th 2020 includes a "first of a set of measures announced in the 2020 European strategy for data. The instrument aims to foster the availability of data for use by increasing trust in data intermediaries and by strengthening data-sharing mechanisms across the EU".

The DGA is a first step of measures to implement PILLAR I of the EU Data Strategy (data voluntarily made available by data holders) with the final objective to make such data easier to access and re-use – FAIR -, to build trust in data sharing and to ensure interoperability among sectors.

A subsequent Digital Market Act (data held by online platforms and originated from business and individual users) is aiming at identifying proper data access and data portability options; while a High-Value Open Government Act in 2021 will mostly be directed towards Public Sector to make such data available for reuse free of charge.

Finally, towards the end of 2021, a Business Data Act in the third quarter of 2021 will address co-generated IoT data from industry and individuals with the final aim to make sure that Europeans stay in control over their data Europeans stay in control over their data vis-à-vis third country jurisdictions.

[8] https://ec.europa.eu/digital-single-market/en/news/proposal-regulation-european-data-governance-data-governance-act

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

In analogy with the main EU Data Strategy, also the DGA is based on four interconnected pillars on top of a common horizontal framework for a Single Data Market: Pillar I addresses the re-use of sensitive data held by public bodies; Pillar II envisages a new role for Data Intermediaries to support voluntary B2B and B2C data sharing (they cannot monetise on the data and should be established in Europe); Pillar III introduces the concept of Data Altruism to meet objectives of general interest under a full transparent mechanism; Pillar IV envisages the creation of a European Data Innovation Board as an expert group advising and assisting the EC in developing a consistent practice for re-use.

The impact of the DGA on the Manufacturing business will be briefly described in section §3.2.1 when we will address the Open Data business model for Manufacturing.

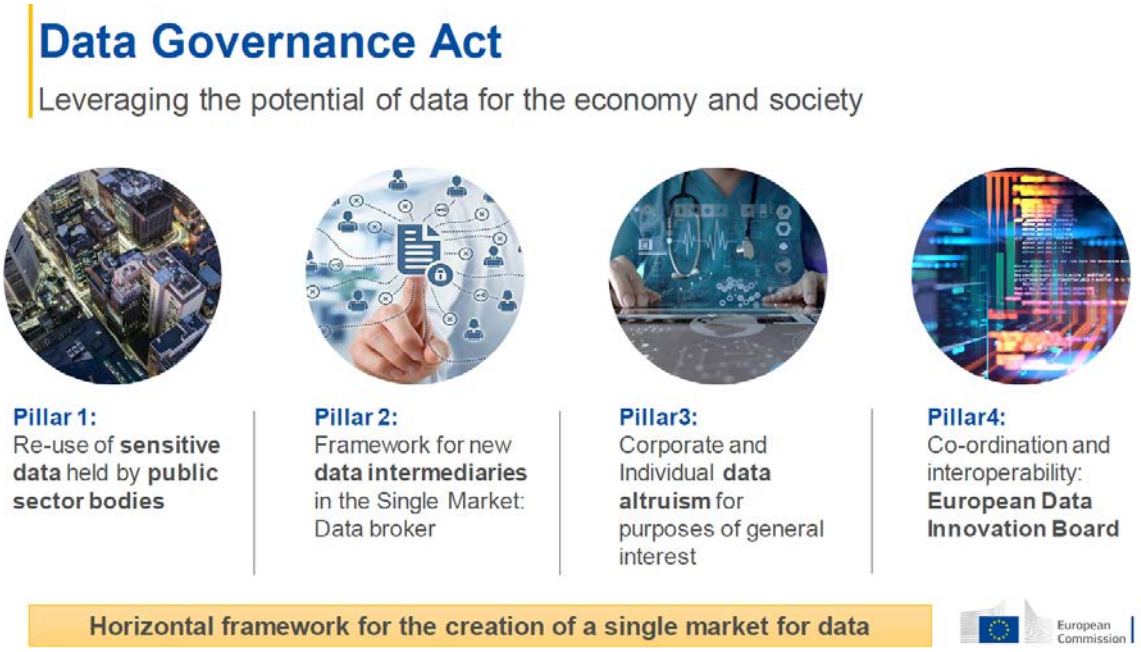


Figure 8 - Pillars of the EU Data Governance Act of November 2020

2.1.2 Roles, Professions and Skills and Data for Manufacturing

A complete analysis of this aspect will be provided in D4.1. In this D2.2 context is more important to underline that the Data Revolution in Manufacturing Industry cannot be implemented without a deep re-thinking of humans' roles and professions, characterised by new pools of distinctive competencies. Related to roles and professions in Industry 4.0, a structured repository of 100+ technical and managerial skills 4.0 covering five areas[9] necessary to define Industry 4.0 strategies, and design, manage and enable Industry 4.0 processes and business models have been developed. In the case of "**Data science management**", five main job profiles and their related competencies validation are defined, as depicted in the following figure.

Data Science Manager. Data Science Managers propose, plan, and manage functional and technical evolutions of the data science operations within the relevant domain. Top related skills for this position are knowledge about data processes, about business processes, about performance indicators, communication with domain experts, ability to manage the data science team and resources, ability to develop and execute the data strategy according to business objectives.

[9] Osservatorio Industria 4.0 - Politecnico di Milano

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

Data Scientist. Data scientists find, interpret, and merge data sources, manage large amounts of data, ensure consistency of datasets, and create visualizations to aid in understanding data. Build mathematical models, present, and communicate data insights and findings, and recommend ways to apply the data. Top related skills for this position are ability to identify and interpret relevant data sources, ability to use a programming language (e.g. R, Python), knowledge about advanced mathematical and statistical models, use of machine learning, use of Bayes classifier, use of Deep Learning techniques, use of Operational Research methods, use of optimization algorithms, knowledge about domain-specific processes, ability to communicate with domain experts.

Data Science Architect. Data Science Architect designs and maintains the architecture of data science applications and facilities. Top related skills for this position are ability to integrate data universe, knowledge about big data architectural standards, ability to select software platforms for big data (Hadoop, Data Lake), ability to select hardware platforms for big data (performances, costs, scalability, flexibility)

Data Engineer. Data Engineers build, manage, and maintain data pipelines. Top related skills for this position are ability to develop data models and workflows, ability to use cloud computing, knowledge about data storage and query languages, ability to maintain security, quality, integrity, safety, and availability of data, ability to integrate new data technologies into existing systems

Visual Data Designer. Data Visualization Designers create custom visualizations from complex data sets in a compelling way. Top related skills for this position are ability to create infographics (maps, charts, diagrams), ability to visualize the huge and complex volume of data, ability to understand complex information, ability to develop insightful and engaging data analytics view, ability to develop vector graphics, scientific illustrations, and icons (maps, charts, diagrams), ability to develop interface and interaction to increase user experience, user experience analysis, design, and evaluation

A validation of this method is part of WP4 and recently started inside the Boost4.0 Big Data for Manufacturing Lighthouse project. More details about the prioritisation of the skills and the evaluation of needed / possessed competencies is included ion D4.1.

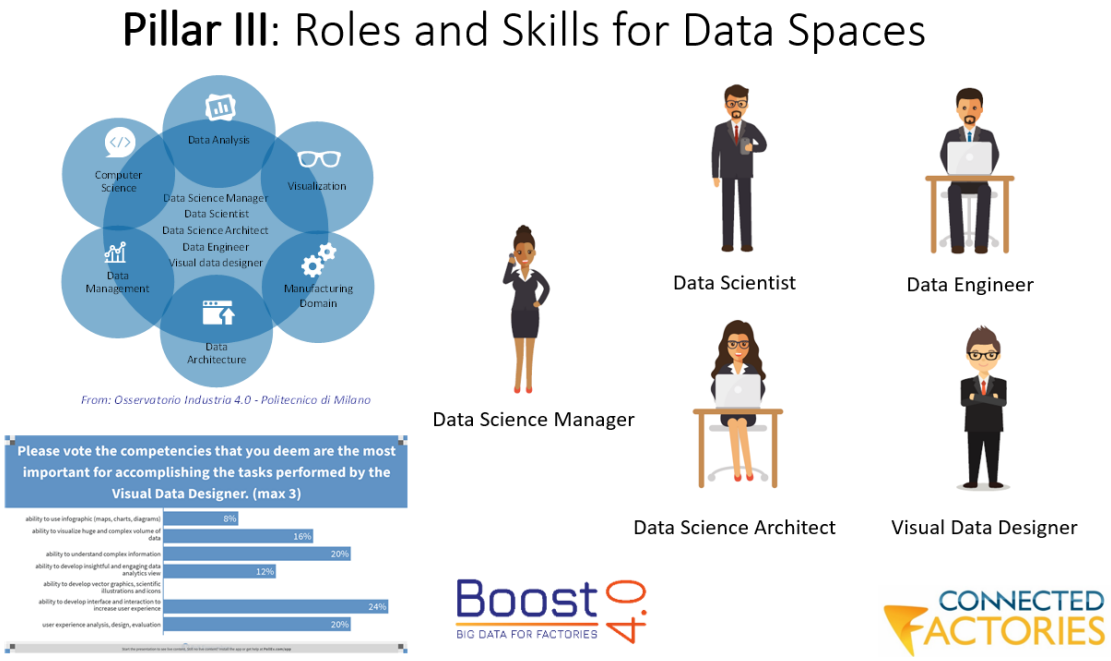


Figure 9 - 5 main job profiles for Data Space management

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

2.2 The AI Revolution and Digital Manufacturing Platforms

The AI, Data and Robotics Partnership[10]'s **Strategic Research, Innovation and Deployment Agenda** (SRIDA) defines the general goals, the main technical and non-technical priorities, and a research and innovation roadmap for the European Public Private Partnership (PPP) on Big Data Value. The activity is driven by the awareness of the advantages brought by a solid European AI system and, indeed, one of the core tasks is to create a European network connecting AI ecosystems.

In the SRIDA it is outlined the context for the operation of the DAIRO Partnership, setting out the European integrated landscape for its adoption and development. The picture below (Figure 1) shows in a schematic way the elements that contribute to define the European AI Framework and Enablers.

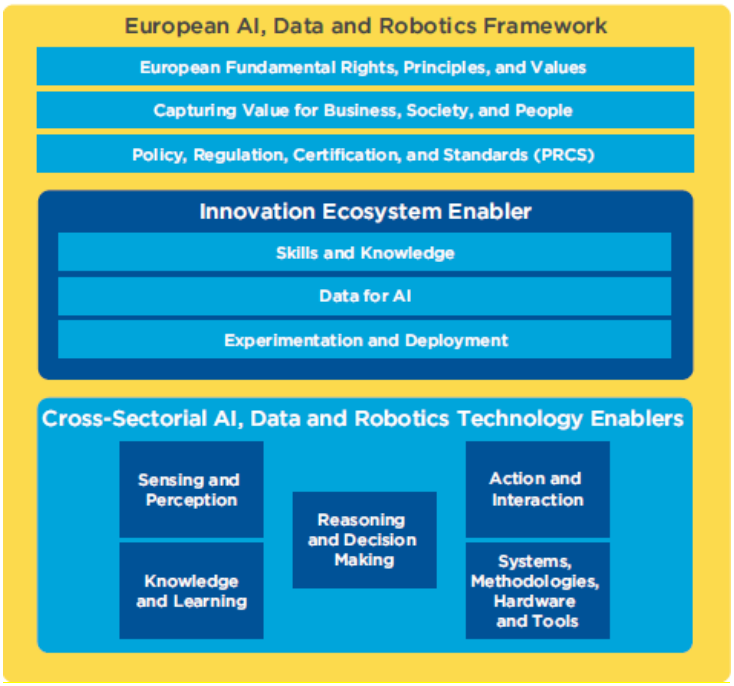


Figure 10 - European AI, Data and Robotics Framework and Enablers released in September 2019

The European AI Framework represents the legal and societal fabric that reinforce the impact of AI on stakeholders and users (see the cluster on the top of the picture) and it embraces several "enablers", that allow its constitution and deployment (cluster in the middle and at the bottom of the picture).

- The AI Innovation Ecosystem Enablers represent essential ingredients for effective innovation and deployment to take place.
- The AI Technology Enablers represent fundamental building blocks required to create the AI system. Originally there were identified five different Technology Enablers (September 2019, see Figure 1); in last release of the document (September 2020) they were reviewed and the list has been updated to six.

According to the last release, each technology enabler is identified by a definition and a list of challenges necessary to put in light the fundamental steps to be overcome in order to implement successful AI products

[10] https://ai-data-robotics-partnership.eu/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

and services. Each challenge is provided with a series of expected outcomes, that can be achieved at **short**, **medium** or **long term**. The six technology enablers identified in last release are:

- **Data Protection and Privacy**: Technologies to preserve the privacy of people, whose personal data are used in modern AI applications. Such technologies have to be deployed in all application involving personal data and it is a matter of enormous importance.
- **Sensing and Perception**: Technologies to create digital information from physical processes and methods to access and assess real-world inputs. It represents the crossover point between the physical world and its digital representation, since it is responsible for collecting all sorts of data and enable their transformation into useful information to support decision-making.
- **Knowledge and Learning**: Technologies and methods that make the sensing, measurement and perception data assets amenable to use in decision-making and able to extract knowledge and insights directly from the data without relying on predetermined expert-defined rules and functions. This involves transforming, cleaning, storing, sharing, modelling, simulation, synthesising and extracting insights.
- **Reasoning and Decision-Making**: The heart of AI, when the output of a model is used to make decisions and take actions. The role may be covered by humans, directly by the machine or by a combination of both.
- **Action and Interaction**: Technologies that embody every aspect of digital and physical AI, working together. For example, interaction may occur between machine and objects, between machines and machines, between people and machines...
- **Systems, Methodology and Hardware**: Technologies that allow data elaboration, from software to hardware and physical structure. They can be of different nature (proprietary or cloud solutions, requiring expertise in the use or more user-friendly), but in any case they are subjected to a strong regulation in terms of privacy, robustness, transparency…

The six technology enablers have been subject of deep analysis in collaboration with almost 20 organisations in the BDV PPP Smart Manufacturing Industry group, in order to apply to the manufacturing landscape the general overview provided in the SRIDA. The goal was to identify which are obstacles and challenges faced in their implementation and future opportunities deriving from their application. The analysis has been conducted taking into consideration three points of view (that is, three different aspects of the manufacturing production):

- The **product**
- The **factory**
- The **supply chain**

To collect opinions from partners and to promote a stimulating discussion where it was possible to take inspiration from others' idea, it has been organized an **interactive online workshop**: for each Technology Enabler, it has been set up a panel containing the challenges identified in the SRIDA, with the possibility to discuss their implication at short, medium and long term.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

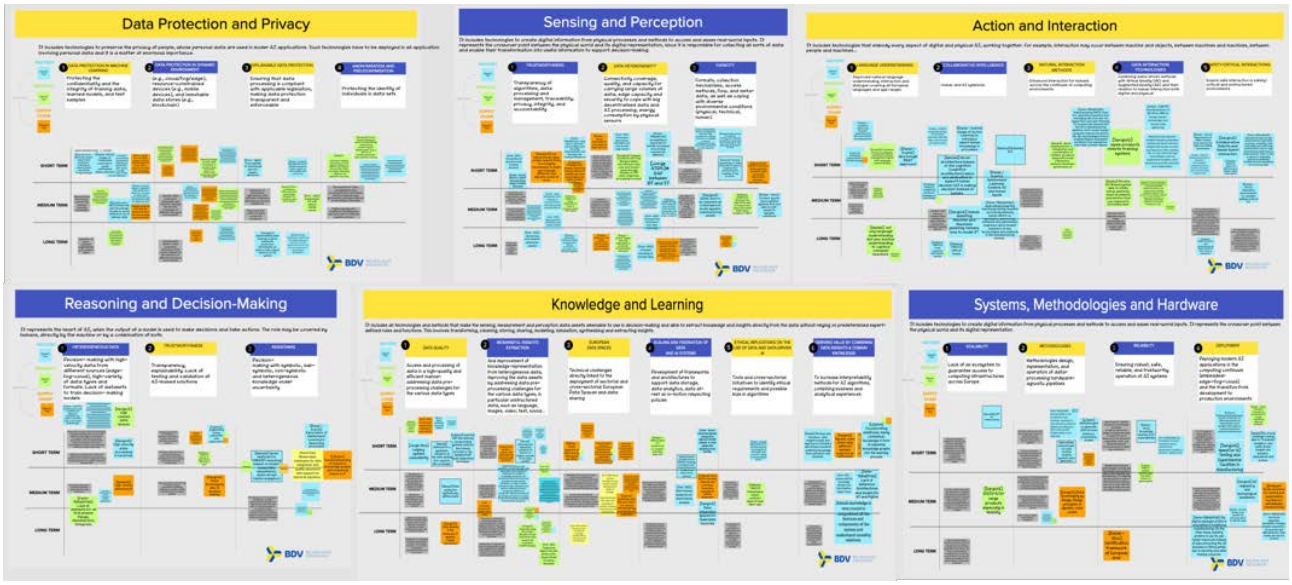


Figure 11 - Six Technology Enabler interactive workshop

At the end of the workshop, more than 200 comments, opinions and reflections have been gathered.

One of the most engaging discussion, emerged from the overall analysis, is about **Standard and Regulation**: the need of defining and adopting a common European semantic standard and of relying on **Reference Architectures** (GAIA-X and RAMI AAS, for instance) came out to be a very important issue to be addressed in the context of AI development.

Moreover, a feel spread among participants is that the European manufacturing landscape required urgently to define a **certification and regulation framework** to be applied in order to guarantee trustworthiness and safety and encourage exchange of data, information and solutions. Not surprisingly, moving to the analysis of "Data Protection and Privacy" challenges, **Data usage control** and **data flow tracing** turned out to generate a common interest, both from the factory, product and supply chain points of view.

Focusing on the exploitation of the data itself, several obstacles have been identified and several solutions proposed. The starting point is to increase **Data potential awareness,** not only generally speaking, that is, understanding at high level the business impact that data exploitation can provide, but also considering the specific data available in the company: people designated to work with data must be trained to have competencies required to deeply understand the meaning and prospective of sources used in the models. Concerning manufacturing domain, the **lack of expertise** related to AI models and Big data is wide and in particular, what it is definitely required is a professional figure skilled both in data and business logics (**domain knowledge**).

Besides that, efforts and resources must be invested for **infrastructures** to store and elaborate data, for **data protection** and to assure **quality** of the data (guaranteed for example by accuracy in input generation, by implementing a suitable pre-processing activity during the ETL phase, by increasing variety and type heterogeneity, leveraging also on **synthetic data**).

Among innovative solutions proposed, it is worth to mention **knowledge graphs**, **cobots** (and **collaborative intelligence**), **5G** and **blockchain.**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

2.3 The Digital Factory Alliance for Digital Manufacturing Platforms

The **Digital Factory Alliance (DFA)** (www.digitalfactoryalliance.eu), is an international trusted community for digital factories stakeholders to foster knowledge sharing and industrial collaboration to achieve data-driven digital transformation.



The DFA has been set-up in 2020 by Innovalia, AtoS Research and Innovation and Engineering Ingegneria Informatica S.p.A as a result of the best practices and lessons learned during the implementation of two of the largest European partnerships set by digital and manufacturing lighthouse industries to trial at scale big data-driven solutions for Industry 4.0 process optimisation (Boost 4.0, Qu4lity).

Industry 4.0 lacks from a common global knowledge platform that facilitates the community a more seamless pilot purgatory experience, collaborating and learning from and with each other and the best and accelerating the digital transformation leverage.

Through adoption of shared models and pilot 4.0 practices, in the medium and long-term, individual piloting activities will result in collective benefits such as better alignment between digital infrastructures and digital manufacturing platforms, more cost-effective system and factory integration and a much larger economies of scale for big data driven manufacturing 4.0 processes implementation. The DFA mission is to take care that such community is nurtured, that such common foundations for data-driven factory transformations are maintained and that successful big data pilots can leverage higher profits to industry in much shorter time scales.

2.3.1 Excellence in data-driven factory

DFA acknowledges the individuality of each factory and the need to develop unique competitive advantages. Yet, to profit globally from data-driven transformations faster and at lower costs, factories should not work alone. Precisely, DFA firmly believes on the need of such individuality to define a set of shared principles that can leverage excellence in data-driven transformation, faster digital value leverage and a successful implementation of resilient and sustainable global production value chains.

The Digital Factory Alliance (DFA) envisions that the successful implementation of outcome-based, resilient and sustainable production value chains depends on the capability of manufacturing factories to embrace true collaboration. **Therefore, the DFA envisions the development of an open and vibrant global business innovation community capable of shaping, steering and harmonising digital technology adoption** (DFA manifesto) **to support future strategic value chains**. The set of **10 principles for excellence** in data-driven factory transformation are collected in the DFA manifesto[11]:

DFactory Manifesto

[11] https://digitalfactoryalliance.eu/wp-content/uploads/2020/10/DFactory-EU-Manifesto.pdf

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

1. Digital is **everywhere**, every factory is **unique.**
2. The digital offer is **infinite** out there, but don't let this overwhelm you.
3. Interoperability by design and **standards** instead of lock in
4. Define your **goals and resources** first and let digital work for you
5. Build your digital **future** with the providers you **trust**
6. Develop increased manufacturing autonomy levels through **experiments** and **evidence**
7. Develop your digital manufacturing capabilities at the **speed of business**
8. Let data and **human-centric AI** drive you to profit
9. Master the **value of zero**
10. Digital **for all**

The Digital Factory Alliance is born to put digital solutions (data-driven AI-powered) to work for advanced manufacturing processes. It should empower manufacturing industry with a friendly, coordinated and low-cost digitisation approach based on robust yet seamless integrations of digital manufacturing solutions and platforms of trust. The DFA mission is to accelerate the sovereign and gradual deployment of resilient, sustainable (green) and outcome-based manufacturing operations and industrial value chains.

2.3.2 DFA Goals

Hence excellence in data driven factory transformation is not a simply a digital technology procurement exercise. Excellence in data-driven transformation demands access to an open community where factories, manufacturing and digital industries can see and can be seen, where they get to know and are known. Hence through open collaboration, the DFA goal is to

- Build **win-win** data-driven transformation scenarios for **digital and manufacturing industries** (future digital value chains).
- **Accelerate** adoption of digital technologies.
- Enable **budget-friendly** deployments.
- Facilitate a balanced digital transformation at the **speed of business growth.**
- Leverage **freedom of choice** for best-in-class and best-in-value digital solutions

To support the digital transformation of manufacturing factories in the realisation of future strategic industrial value chains, the DFA will leverage the following 5 Transformation Assets:



1- Marketplace

- The DFA let factories be the first to adopt.



2- Certification Framework

- The DFA let factories be the first to trust.



3- Innovation Campus

- The DFA let factories be the first to know and inspired



4- Experimental Facilities

- The DFA let factories be the first to try.



5- Digital Factory Network

- The DFA let factories be the first to profit.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

2.3.3 DFA 2021 Plan

The main objective in 2021 is the ramp-up and consolidation of the Digital Factory Alliance. The main goals for the DFA in 2021 should be:

- **Goal 1 – COMMUNITY:** Establish and grow the DFA community (digital and manufacturing champions, strategic value chains and DFA transformation assets).
- **Goal 2 – COLLABORATION:** Gain DFA community visibility and attractiveness through aggregations and articulation of a compelling programme of DFA activities.
- **Goal 3 – CONNECTION:** Design a solid DFA collaboration framework through partnerships and agreements, avoiding isolation and fostering engagement and cooperation with main digital transformation initiatives., communities and networks.

In order to achieve these 3 Goals, the DFA will kick-off 2 lighthouse Initiatives:



Resilient Leadership in Manufacturing Repurposing Networks



Zero Blackbelt Digital Factory Network

Additionally, the DFA will perform a number of activities for:

- Networking events
  - Deep Factory Trials
  - Meet the Digital Factory Network
- Digital Tech talks
  - "Digital Platform Innovation for Zero Blackbelts"
  - "Manufacturing Repurposing COVID-19 Products"
- "Know the DFA" webinars
  - Know the DFA Marketplace
  - Know the DFA Testbeds & TEFs
  - Know the DFA Data-Driven Reference Architecture

In 2021 it is also planned to set up the European Network of Digital Factories, collecting the most advanced data driven large scale trials in Industry 4.0. This initiative will be pursued in close collaboration with the WEF Global Network of Lighthouse Factories (https://www.weforum.org/projects/global_lighthouse_network), bringing forward the best examples of digital transformation driven by big data and AI.



European Network of Digital Factories

DFactory

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

3 Business aspects towards a European Data Space for Manufacturing

The development of a common EU Data Space for Manufacturing implies several considerations from a business point of view. In §3.1 the overall market perspective is described, making reference to the EU Data Strategy itself and a couple of studies addressing the topic of the data-powered enterprise and of data accessibility for AI applications. In §3.2, leveraging on the April 2018 EC communication "Towards a common EU Data Space"[12] and its Guidelines for the Private sector[13], a detailed analysis will be performed to identify interesting Data Economy new business models for the manufacturing sector.

3.1 The Market Perspective: from data-powered enterprises to Data4AI perspective

This section will present the results of two market studies about the business impact of Data Spaces. The former is focussing on the so-called data-powered enterprise, i.e. an enterprise which is basing its decisions on data. The latter is more related to the AI revolution and to what extent it depends on high quality and high value Data4AI assets.

3.1.1 The Data-powered Enterprise (CapGemini)

The EU data strategy has identified some global trends which are substantially affecting the Manufacturing Industry as well: in the 2018-2025 period, global data volumes will dramatically grow from 33 zettabytes to 175 zettabytes (5x) and the data processing infrastructure will radically change from 80% centralised (and 20% distributed) to exactly the opposite, where smart connected objects will process the 80% of the available data. *In this perspective the so-called data computational continuum from the manufacturing shop floor till to the global cloud will make the far edge architectural level even more important and relevant for Factories of the Future. New architectures and solutions for highly distributed data storage, computation and sharing need to be experimented and validated in our enterprises. This will be more deeply analysed in Chapter 4.*

Moreover, on a more business perspective, in the same seven years the value of the data economy will be multiplied by 2.8 reaching the amount of 5.8% of EU GDP; number of data professionals in EU27 will double reaching the value of more than 10 million people, while the percentage of EU population with basic digital skills will reach the 65%. *This poses new challenges for our society but also for our business entities: while at today the most important and competitive manufacturing Countries in EU (Germany Italy France) are not in the first positions of the DESI index (Digital Economy and Society Index), in the near future digitalisation and data orientation will become a must for enterprises to remain competitive on the global market, also as a consequence of recovering from the current pandemic crisis*.

In its report "The data-powered enterprise: why organizations must strengthen their data mastery"[14], *CapGemini conducted a deep analysis on how companies base their critical decisions on data. They surveyed 500 technology executives and 500 business executives, conducted detailed interviews with more than 15 executives and looked at the relationship between data mastery and the financial performance of organizations. They found that organizations still have some ground to cover when it comes to data-driven decision making and actioning, with only 39% successfully turning data-driven insights into a sustained competitive advantage. This may be because only 20% of business executives trust their data, often because*

[12] https://ec.europa.eu/digital-single-market/en/news/communication-towards-common-european-data-space
[13] https://ec.europa.eu/digital-single-market/en/news/staff-working-document-guidance-sharing-private-sector-data-european-data-economy
[14] The data-powered enterprise (capgemini.com)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

it is of poor quality. The small minority that do, however, enjoy a significant performance advantage over the rest, seeing up to 22% higher profitability and 70% higher revenue per employee. These are the "Data Masters."

In particular, regarding the data-driven decision making critical processes, the situation of the manufacturing industry, in comparison with other industries, is quite worrying with just the 43% of cases with respect of more than 65% for banking and 55% for insurance and telco.

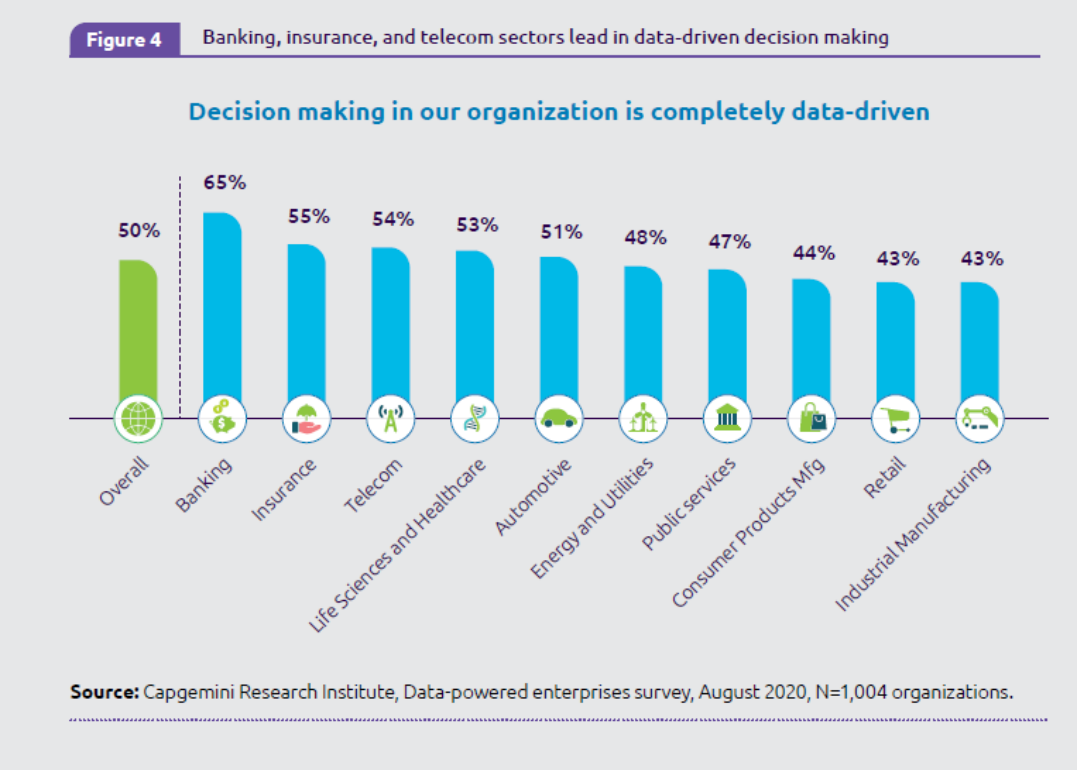


Figure 12 – How many enterprises make Data-driven decision, by sector (The data-powered enterprise, August 2020)

In the same report, CapGemini depicts an evolutionary pathway for enterprises to achieve the status of Data Masters. In a classical bi-dimensional quadrant diagram, they put on the X axis the technological aspects for data driven decision making (Data Foundations) and on the Y axis the organisational, skills soft aspects (Data Behaviours). Four categories of enterprises (out of a sample of 1000+ companies) emerge: Data Laggards, Data Enabled, Data Aware and Data Masters. Their typical attitude towards Data Innovation is reported in a table with nine analysis dimensions.

For instance, in the **Technology for Data** dimension, Data Laggards face Siloed data sources, for Data Enabled integrated data platforms do exist but alongside legacy platforms, which makes data discovery and exploration confusing; for Data Aware companies Integrated data platforms do exist but querying and reporting takes time. For Data Masters, Unified data platforms leveraging cloud and self-service tools are provided.

In the **People** dimension, in Data Laggards enterprises Technical teams handle business requirements, in Data Enabled the business is dependent on the technical teams; for Data Aware companies each BU has dedicated

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

resource(s) that can run simple queries and extract reports. For Data Masters, business users are empowered to build their own reports and dashboards.

In the **Data Monetisation** dimension, Data Laggards do not see any opportunities to monetize data, in Data Enabled enterprises monetization is not seen as an opportunity; for Data Aware companies business is keen on monetizing however it is not supported. For Data Masters, Data is an integral part of decision making and can even monetize data.

Cohort	Data laggards	Data enabled	Data aware	Data masters
Features	Basic reporting capabilities and minimal visualization	Data and the infra for reporting and visualization exists but no clear capabilities on effectively using this for business	Business is keen on leveraging data and insights but are handicapped due to issues of data quality or skills or governance	Advanced reporting and analytical capabilities exist in each business area
Data discovery	Challenging for business users and new staff to understand the data	Data discovery is complex due to restricting policies around access and usage	Data discovery is complex due to lack of strong master data management	Well maintained master data and data catalogues allowing easy data discovery
Relationship between technical and business users	Business depends on technical teams for all reporting needs	Data is available for analysis, but business is not able to leverage this for insights	Business users are aware of the kind of analyses they need but they are not easy to obtain	Business and technical teams collaborate to understand how they can reach business objectives
People	Technical teams handle business requirements	Despite the availability of the required tools and platforms, the business is dependent on the technical teams	Each BU has dedicated resource(s) that can run simple queries and extract reports	Business users are empowered to build their own reports and dashboards
Training	Minimal focus on upskilling technical or business users	Training on data and analytics is available to technical users but minimal training exists for business users	Training programs exist for selected business users	Continuous focus on upskilling majority of the business users
Processes	Minimal guidelines on data access, privacy and none on ethics	Guidelines on data access, use, and privacy exist, but limit the ability of business to explore data	Guidelines on data access, use, privacy, security exist but no standards defined for ethics	Robust framework covering all aspects of privacy, security, ethics apart from access and use
Technology for data activation	Siloed data sources	Integrated data platforms exist alongside legacy platforms, which makes data discovery and exploration confusing	Integrated data platforms exist but querying and reporting takes time	Unified data platforms leveraging cloud and self-service tools
Data advantage (Use of external data sources)	No thinking on how to leverage external data	Data platforms make the external data sources available for analysis, but cross-analysis is not facilitated	Business looks at external data but integrating with other types of data is still complex	Actively work with external players to collect various sorts of data and leverage this for insights
Data monetization	No opportunities to monetize data	Monetization is not seen as an opportunity	Business is keen on monetizing however it is not supported	Data is an integral part of decision making and can even monetize data

Figure 13 – CapGemini evolutionary Data pathway for enterprises (The data-powered enterprise, August 2020)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

In the four-quadrants diagram, the 1000+ enterprise have been positioned regarding the nine variables above and relevant questions. With no surprise, Data Laggards count for the vast majority of the cases (71%) and Data Masters for just 16%.

However, it is quite significant that Data Aware and Data Enabled cases are underrepresented. The interpretation of this polarisation can be that the pathway to Data Mastering usually does not separate Technology (Data Enabled) from Behavioural (Data Aware) aspects, so that the data driven digital transformation process mandatorily should include both aspects in methods and tools which simultaneously are able to stimulate and develop data technologies and data-oriented processes and skills.

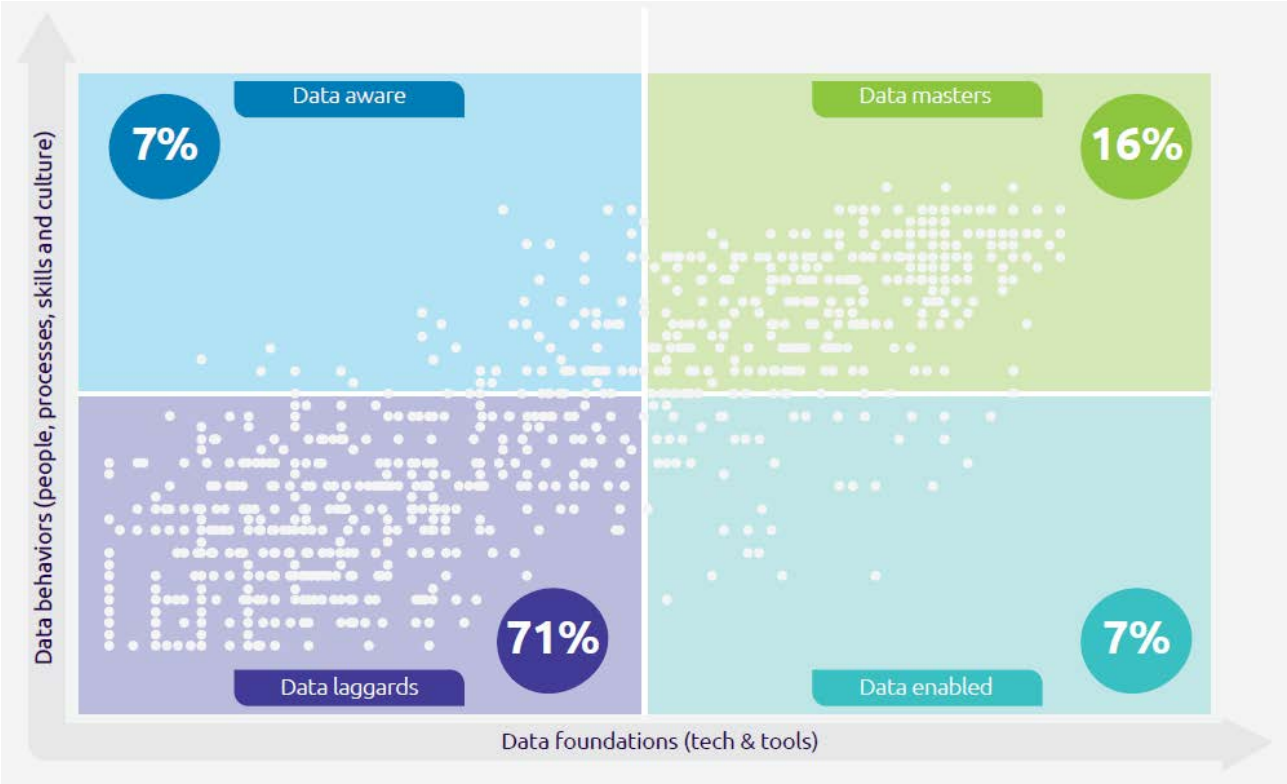


Figure 14 - Capgemini four classes of Data exploitation for enterprises (The data-powered enterprise, August 2020)

Finally, the report elaborates its own recipe to become a Data Master. Organizations should focus on three priorities:

- **Data activation (on the Y axis)**: this refers to the use of data in end-to-end business processes to secure business outcomes i.e., transforming data and insights into actions. To achieve this, organizations need to: align their data and analytics strategy with the business strategy, build business users' trust in data, establish an AI and analytics co-existence to assist business teams, foster a data-powered culture by strengthening "data citizenship" or "data culture".
- **Data enablers (on the X axis)**: these are the key capabilities necessary for building data-powered enterprises and include infrastructure, governance, and operations. Organizations need to strengthen data collection processes and improve data quality, invest in data landscape modernization to get agility in data activation, operationalize data and analytics through DataOps and MLOps, adapt their data governance as data mastery evolves.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

- **Data advantage (on the diagonal)**: this is about creating a competitive advantage, particularly by leveraging external data, including data from hyper-scalers. To realize this, organizations need to be able to leverage external data to enhance their insights.

3.1.2 The Data for AI Challenge (Sloan MIT Review)

In the article **The Data Problem Stalling AI**, authors strongly affirm that AI efforts can fail to move out of the lab if organizations don't carefully manage access to data throughout the development and production life cycle [15]. The point of Data Quality is crucial: Data quality encompasses four dimensions: intrinsic (accuracy), contextual (completeness), representational (interpretability), and accessibility (availability or easily and quickly retrievable). To be considered high quality, data must satisfy all four dimensions.

Five main phases are identified in the Data2AI migration process in an enterprise as well as the relevant stakeholders, the locus of data and the focus on Data Accessibility efforts.

Phase 1: Ideation. The ideation phase serves as a filter to identify potential high-level business cases for AI in the organization. Most conversations during this phase are between managers, business consultants, and scientific advisers (who are sometimes also full-time academics). The goal is to create a meeting space for business and science. The resulting business cases should look promising and feasible. In AI consulting companies, this crucial first step serves to educate clients on the potential of AI. During this phase, however, the emphasis is on data existence rather than data accessibility. Discussions revolve around business objectives and the application of AI models to address the organization's current problems.

Phase 2: Blueprint. Not all use cases generated during the ideation phase will be selected for implementation within a given period because of priorities, resource constraints, or a lack of potential value. During the blueprint phase, a comprehensive use case is generated. This includes details such as clear and measurable business objectives, an action plan that outlines specific AI techniques, and the data elements that should be available to feed AI. During the blueprint phase, data accessibility is still assessed solely on the existence of data, because sights are set on the next phase of the process, which is to build a working prototype. The underlying assumption is that if the data is there, that's good enough, because it allows the team to move forward.

Phase 3: Proof of concept. During the proof-of-concept phase, data scientists build one or more models to implement the agreed-upon use cases. Most of the work is focused on iteratively creating, training, and testing models to measure their relative performance against one another and to see whether AI actually lives up to expectations with new input. Data is extracted from source systems and transformed by data engineers so that it complies with the format and accuracy requirements of the models under construction. Although the solution may ultimately be delivered through an application with a user interface or tightly integrated within the organization's business processes (to alter a credit application process in a bank, for example), the proof-of-concept phase typically does not focus on those efforts just yet. Similarly, teams focus on getting the data to advance their work in the short term, giving little consideration to how data will eventually be accessed once the AI goes into production.

[15] https://sloanreview.mit.edu/article/the-data-problem-stalling-ai/ SLOAN MIT REVIEW

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

Phase 4: Minimum viable product. Once a variant of the proof of concept demonstrates sufficient value, it is refined into a minimum viable product, or MVP. At this point, data scientists and data engineers step back and software engineers take over, given that the AI will eventually leave the lab, be deployed within the organization's infrastructure, and get integrated with other production systems, if applicable. An unintended consequence of the strong focus on model development in the previous phases is that considerations regarding the accessibility of data in production have taken a back seat. Once software engineers and IT staff become more involved in discussions about the specifications and the integration of the solution to be delivered, questions related to data accessibility may reveal that a crucial feature used by a model requires significant, unplanned work.

Phase 5: Production. In this last phase, the refined MVP that contains the AI is released into production and must now be fed with data sourced directly from production systems. Data may need to be pulled from multiple systems and transformed to generate the required input for the model to support the business case in production. Whether this happens in real time or in batches (for example, to retrain and retest a model at frequent intervals), this is where the real issues related to AI integration emerge, especially with respect to the organization's data infrastructure. If data cannot be provided, extracted, and integrated by autonomous systems at the required volume or velocity (due to legacy systems, for instance), the AI may lose all of its potential value.

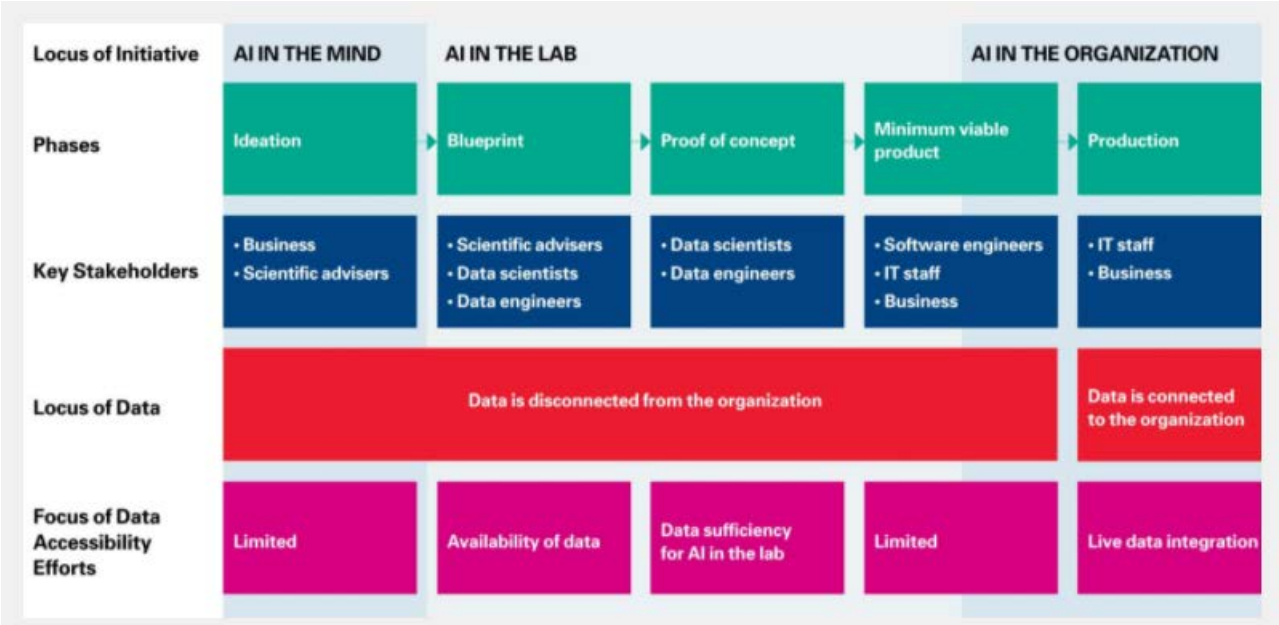


Figure 15 - AI exploitation framework, Sloan MIT Review

Authors finally identify four main misconceptions about Data Accessibility for AI. In addition to understanding the different roles and phases of AI development and their impact on data accessibility, it is helpful to understand some basic misconceptions about the nature of data and how it is perceived in many organizations.

Misconception No. 1: Data accessibility is a technical issue. Technology problems, while often complex, can usually be fixed with the right talent and resources. Participants in our research argued that data accessibility is really a management issue that involves technology. AI solutions must start with a clear understanding that complete, accurate, and timely data has no value if it cannot be retrieved quickly and easily. The fact that

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

data is located somewhere across multitudes of databases and spreadsheets does not necessarily mean that it is accessible. Sometimes data accessibility issues exist because data governance or security policies restrict access.

Misconception No. 2: Data is merely a by-product of operations. This misconception is often seen in organizations where analytics and AI efforts sit apart from operations — and where AI's potential to improve or revolutionize processes across the organization has not yet been recognized. As a result, operational systems (such as enterprise resource planning and customer relationship management) consume and produce data, but there is no understanding of the potential value of this data for AI. If analytics or AI teams want to use data from operations, they have to retrieve and leverage it on their own, similar to what traditional data warehouse teams have done for many years.1

Misconception No. 3: Data accessibility can be addressed in the later phases of the AI life cycle. The five phases of the AI life cycle are designed to push AI teams to work in an agile mode, especially during the proof-of-concept and MVP phases. The very nature of AI as an uncertain endeavour lends itself well to this approach. Teams must be able to experiment with models and pivot on emergent results to find the optimal solution to the organization's problem. Unfortunately, this also encourages teams to focus almost exclusively on the scientific portion of AI work for the better part of the first three phases. The stakeholders involved during the ideation, blueprint, and proof-of-concept phases are not the ones who deal with data accessibility issues. Data engineers are primarily concerned with creating flat files that data scientists can use to build and train models, and any means within their reach to generate those files — including hacks, work-arounds, and simulated data — is considered fair game.

Misconception No. 4: Data in the lab and data in operations are the same. Companies are becoming highly skilled at building AI-enabled proofs of concept. However, the real test is whether they can move past the controlled lab environments of the proof-of-concept phase to the messy production environments. Often, the assumption is that the data retrieval process for the proof-of-concept phase can be replicated at little to no cost once the AI moves through MVP and then into production. But recall that data in the proof of concept comes from a few flat data files that were specifically created — often from historical data snapshots — for the purpose of building and testing models.

On the basis of the above theoretic considerations, inside WP2 we have ideated and developed the DS pathway as reported in Chapter 5.

3.2 Business Models for Data Spaces in the Private Sector

As a guideline (commission staff working document) accompanying the communication "Towards a common European data space", the EC released on April 25th 2018 the document "Guidance on sharing private sector data in the European data economy". For the private sector, the EC identifies three main models to B2B data exchange:

a) **An Open Data approach**: The data in question are made available by the data supplier to an in principle open range of (re-)users with as few restrictions as possible and against either no or very limited remuneration.

b) **Data monetization on a data marketplace**: Data monetization or trading can take place through a data marketplace as an intermediary on the basis of bilateral contracts against remuneration. Suitable when either

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

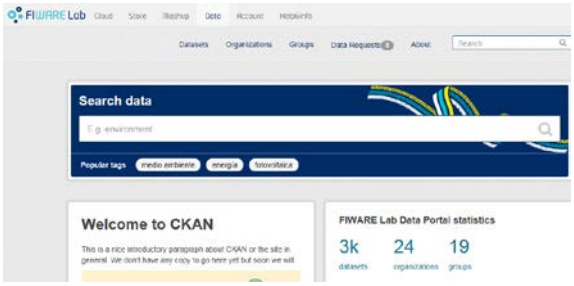
(1) there are limited risks of illicit use of the data in question, (2) the data supplier has grounds to trusts the (re-)user, or (3) the data supplier has technical mechanisms to prevent or identify illicit use.

c) **Data exchange in a closed platform**: Data exchange may take place in a closed platform, either set up by one core player in a data sharing environment or by an independent intermediary. The data in this case may be supplied against monetary remuneration or against added-value services, provided e.g. inside the platform.

In the Manufacturing sector, we can identify different instantiations of such generic models as indica ted in the three paragraphs below §3.2-1-2-3.

3.2.1 Open Data in Manufacturing

By nature, manufacturing enterprises base their global competitiveness on proprietary data and confidential knowledge which should be protected along the lifetime of the company. What is it therefore the role of Open Data in Manufacturing? In the following we will mention two of them but especially in the context of H2020 and future 2021-27 MFFs we believe these models should be much more developed and pursued. The first example stems upon the concept of Didactic Factory (or Learning or Teaching Factory) which is the typical Testing and Experimental Facility where to conduct "test before invest" proofs of concept maybe inside a Competence Center or a Digital Innovation Hub. Data produced by a pan-EU network of (AI) TEF for Manufacturing can become a fantastic Open Data Space for ecosystems of (young) experimenters, start-ups and SMEs. In the DIMOFAC and in the AI REGIO I4MS project, such a network is made operational, by a federation of e.g. SmartFactory KL in Germany, Swiss Smart Factory, MADE Competence Center in Italy, Brainport Innovation Center in the Netherlands, Automotive Innovation Center in Spain and many others. We have identified some steps to be followed in order to implement such a network, inspired for instance by ODINE (Open Data Incubator Europe) or the FIWARE Data Lab (see picture below).

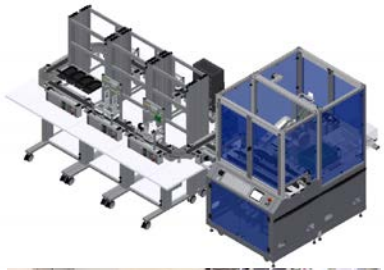


The Dataverse® Project

ODINE Open Data Incubator Europe

SWISS SMART FACTORY ®





INDUSTRY4.0 lab

innovalia ASSOCIATION

DFKI | smartFactory KL®



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

a) **Open and Standard Data Models** such as the RAMI AAS standard should be shared in the network, so that to model Smart Industry entities (such as robots, AGVs, machines, conveyors) and provide the static part for Manufacturing Digital Twins. Data integration facilities are able to unify data models in a Context and for a Scope, integrating Data in Motion from Industrial IoT and Data at Rest from other data sources. The EIT Manufacturing project CanvAAS develops tools to allow industrials to build data models based on the AAS standard, and deploy them to the Eclipse BaSyX industrial middleware for AAS.[16]

b) **Network of open Didactic Factories** producing and sharing their data thanks to standard protocols and data formats (e.g. OPC-UA, MQTT, ROS, AMQP). Networking so called didactic factories is just in its infancy, but one major result of this integration effort is the SmartFactory.eu[17] initiative recently launched by three main teaching factories in Germany, Belgium and the Netherlands.

c) **Data Transformation techniques and Synthetic Datasets**. In the presence of non-public Data to be shared, new (AI-based) techniques such as Aggregation, Filtering, Anonymization, Pseudonymization can be applied to raw data and make them compatible with the Open Data approach. Moreover, big expectations are supposed to come from next generation Data Emulators and their synthetic datasets.

d) **One stop shop**. A common Data Space for the pan-EU network of didactic factories is able to combine centralized metadata catalogues and data ontologies with highly distributed data storage spaces (iSpaces and EUHubs4Data spaces) all of them sharing search-browse-discovery-selection primitives in one-stop-shop

e) **Ecosystem of Innovators.** This final step is aimed at involving the network of DIH4Industry Hubs in order to engage and involve a critical mass of innovators (startups and SMEs) able to populate the pan-EU Didactic Data Space with new innovative (AI based) on top of the open data space (Data-AI Community)

A second typical example of interoperability between open and closed data spaces comes from the **Urban Manufacturing** concept, a **Smart Factory in a Smart City**, where manufacturing enterprises are expected to exchange and share their data with the surrounding urban community and vice-versa. Typical scenarios of Urban Manufacturing UM regard not just **Garage UM** (do-it-yourself FabLabs and 3D printing shop floors) or **Home UM** (small labs especially for the clothing and food industry) or **ZeroKM UM** (very local value chains) or other new forms of manufacturing. In the more traditional manufacturing sense, a **Twin UM** scenario realizes when factory and city grew together in a harmonic and symbiotic manner like for instance a 60 years old factory in the suburbs which has been incorporated by the extraordinary growth of the city. Finally, **New-borne UM** scenarios could occur in case of re-shoring or re-location of a zero-impact, zero-emissions factory with extraordinary less environmental impact when implemented in an Urban environment. Finally, another form of UM is the **Proxy UM**, where local dimension of the factory, its processes and its employees is one critical success factor in this global economy (and COVID recovery economy).

Typical examples of interaction between Smart Factory and Smart City regard the inbound and outbound logistics (especially when moving large size components or products), the optimisation of employees' mobility (public transport and shared vehicles) and the minimisation of environmental impact regarding both air / water pollution and waste management in general.

[16] EIT Manufacturing, CanvAAS project. https://eitmanufacturing.eu/canvaas-a-common-language-for-machines-entering-the-industry-4-0-era/
[17] https://smartfactory.de/en/smartfactory-eu/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

3.2.2 Data Marketplaces in Manufacturing

Pure Data Marketplaces in the Manufacturing sector have not been developed so much so far. Much more interesting is when data selling is associated to new service-oriented business models and/or to selling of physical goods, such as equipment.

In the NMBP-20 call of 2018, **A digital 'plug and produce' online equipment platform for manufacturing (IA)**, the aim was to (a) design and build the digital platform that brings together suppliers and users in a transparent and efficient way; and (b) populate it with adequate product information. This will constitute a set of pilot implementations intended to sell 'plug and produce' industrial equipment and services to customers globally. The platform should therefore facilitate B2B transactions and host associated services in the form of digital product models allowing users to simulate (e.g. digitally test) the capabilities of the equipment on offer and its compliance to standards. This will ultimately boost product quality, transparency and usability based on Return on Experience.

Two IAs have been funded in this call and represent pioneer initiatives in the domain of data marketplaces for manufacturing. **Weldgalaxy** (Digital Dynamic Knowledge Platform for Welding in Manufacturing Industries) will deliver a B2B online Platform that brings together global buyers (end-users/OEM) and EU sellers (manufacturers/suppliers/distributors/service providers) of **welding equipment** along with auxiliaries-consumables and services, thereby enhancing the visibility of EU's welding products-prototypes-services to global users (via digital marketing strategies) and providing innovative web-based services (e.g. equipment selection and inventory management, digital design/testing of equipment capabilities) to boost EU market share and competitiveness. The digital platform will incorporate Knowledge base engineering (KBE) tool that streamlines equipment selection process for end-users and allows 'plug and produce' digital manufacturing of the right equipment to specified customers'/end-users' requirements and regulatory compliance.

Much more interesting from our perspective is **MARKET4.0** (A Multi-Sided Business Platform for Plug and Produce Industrial Product Service Systems) where the typical functionalities of a Marketplace (catalogues, matchmaking, accounting, e-commerce) are developed on top of a Data Sharing network. The objective of MARKET4.0 is to define, develop and validate an open multi-sided marketplace, based on a trusted P2P data sharing infrastructure for Industry 4.0, that brings together Industrial Product Service Systems (IPSS) providers (supply side) and OEMs as their customers (demand side). It will allow direct interaction among the different sides to improve the sales power of production equipment SMEs. MARKET4.0 will offer advanced web-presence of production equipment SMEs extended by functionalities such as simulations, VR/AR capabilities and a P2P (Industrial Data Space) offering smart user-services and a secure API to try and test the Digital Twin of the production equipment, on top supplier and customer data. This will facilitate direct transactions between market peers (supplier-to-supplier, supplier-to-customer, customer-to-suppliers and more) during the entire B2B lifecycle, i.e. from equipment search to procurement and commissioning. MARKET4.0 will create trust in the business transactions between SME production equipment manufacturers and their customers, as an integral element of the **Industrial Data Space (IDS)** reference architecture. It includes technological trust (simulation before purchasing), financial trust (Blockchain distributed ledger technology for secure payment), delivery and (anonymized) feedback. In MARKET4.0 the supply side (mainly SMEs) may offer a) production equipment, b) services that extend the capabilities of the production equipment, c) production equipment as a service and d) collaborative engineering services. The marketplace will generate additional value by providing engineering services and acting as a mediator between suppliers and customers.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

While Open Data sites are extremely dynamic both in the demand and in the supply side, MARKET4.0 implement a hybrid model where the offer-side of data and services is a typical "trusted data network" (see next paragraph), while the demand side is quite dynamic and agile, exploiting the platform credentials to join on-the-fly the trusted data network. Three main sectors are addresses: the metal working, the plastic and the high-tech manufacturing domains. In particular, the third domain (implemented in South Netherlands) addresses the problem of managing requests for quotation and orders in very dynamic "manufacturing as a service" networks.

The **Smart Connected Supplier Network** (SCSN[18]) is one of the most promising use cases based on IDS components. Digital collaboration in the supply chain is one of the goals of Smart Industry, the Dutch national Industry 4.0 programme. However, it is not easy to link ICT systems between companies, especially covering the supply network as-a-whole. Current practice shows that many companies solve this manually: Employees transfer information from an incoming order into their own system. SCSN is a communication standard enabling the machine building industry to share data across company borders in an easier, safer, and more reliable way. SCSN is led by IDS-member Brainport Industries, sector organization for the Dutch high-tech supply sector and supported by TNO. TNO is an independent research organization in The Netherlands participating in the IDS Launching Coalition. SCSN encompasses a nucleus of a manufacturing data spaces where companies are able to exchange and share data in a secure and trusted way, this way anticipating the DEP initiative about manufacturing data spaces in dynamic value chains.

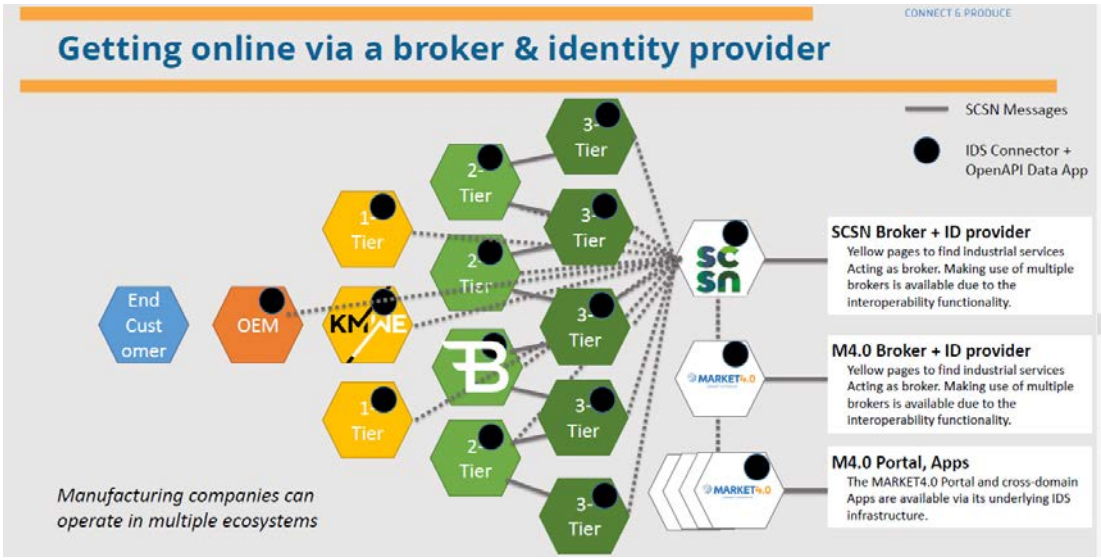


Figure 16 - SCSN structure for manufacturing data exchange

In this respect, just started is the NMBP-40 project **DOME** (Digital Open Marketplace Ecosystem 4.0) aiming at developing a comprehensive industrial data ecosystem aligned with the Open Science and Open Innovation objectives to enable sharing of business-to-business (B2B) data for the purpose of value generation and creation of new or enhanced products, processes, and services. DOME 4.0 will be open to all providers and users of data, and aims to facilitate maximum knowledge extraction with the help of ontology-driven semantic data interoperability and modern data processing technologies adopted from the fields of Machine Learning (ML) and Artificial Intelligence (AI). These features are crucial to scale and advance the

[18] https://www.internationaldataspaces.org/the-smart-connected-supplier-network-by-tno/

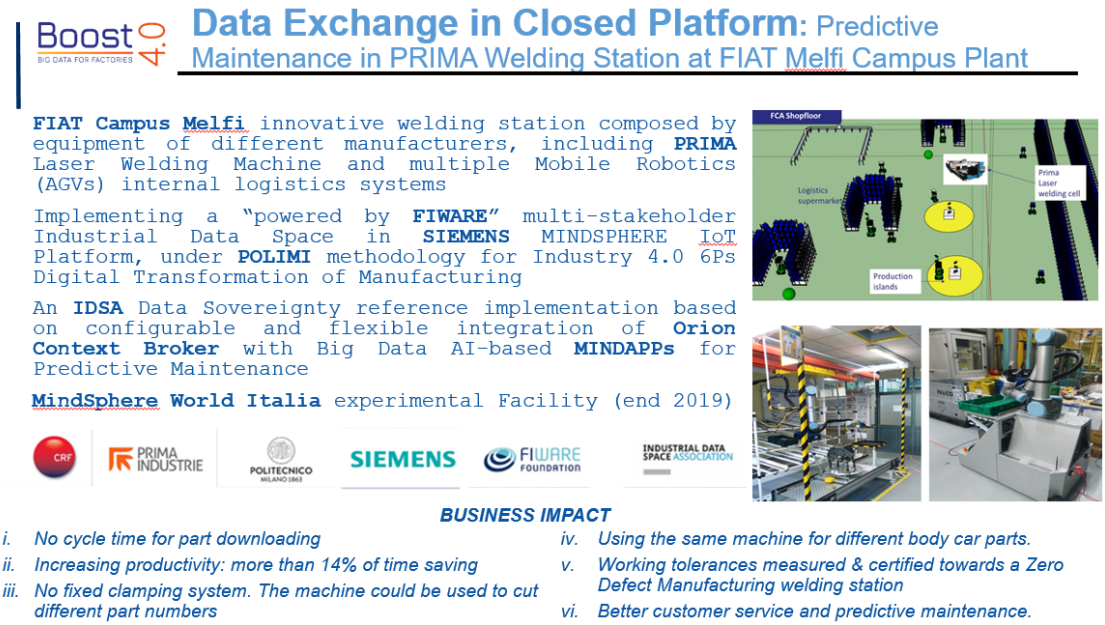
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

proposed ecosystem to any sector of the economy. Given the significant contribution of the materials and manufacturing sectors to the European economy, DOME 4.0 focuses on data-driven knowledge generation within these key sectors.

3.2.3 Trusted Data Networks in Manufacturing

The third and most diffuse business model for implementing Data Spaces in Manufacturing is the Trusted Data Network which is the typical implementation of manufacturing value chains. Such a model does exist since many years and in most cases implemented by a Large Manufacturer (e.g. in Automotive or Aeronautics) which imposes its technology, data models, access and usage rights to the whole hierarchical tier-1 tier-n value chain. The challenge here is to progressively substitute static, rigid and paper agreements (NDAs) with flexible and agile ICT Data Spaces, in order to support non-hierarchical and dynamic business ecosystems, where new actors are continuously joining the network maybe also for a very limited time span or just for one transaction.

One example of how this agility can be implemented is represented by the Boost4.0 project[19]. The FIAT-PRIMA pilot combines the traditional scenario of a large car manufacturer sharing data with a Machinery Manufacturer, to allow the latter to perform fleet management and maintenance operations on its machines. The access and usage rights about the FIAT data (including the data coming from PRIMA laser machine) are regulated by flexible, dynamic, on-the-fly configurable Data Sovereignty models. Moreover, the lighthouse pilot is implemented via the SIEMENS Mindsphere Industrial IoT platform this way bridging the gap between open data models and commercial industrial data platforms. More and more interesting is also the fact that the Predictive Maintenance app is a Mindsapp implemented by a third party (who won the FIAT Melfi Campus hackathon last year) on top of the FIWARE for INDUSTRY open source toolkit. This makes the case particularly interesting from a data-driven business modelling perspective, where the four stakeholders (FIAT, PRIMA, SIEMENS and SME) involved are able to models their business interactions through the formalism of the Business Model Navigator.



[19] https://boost40.eu/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

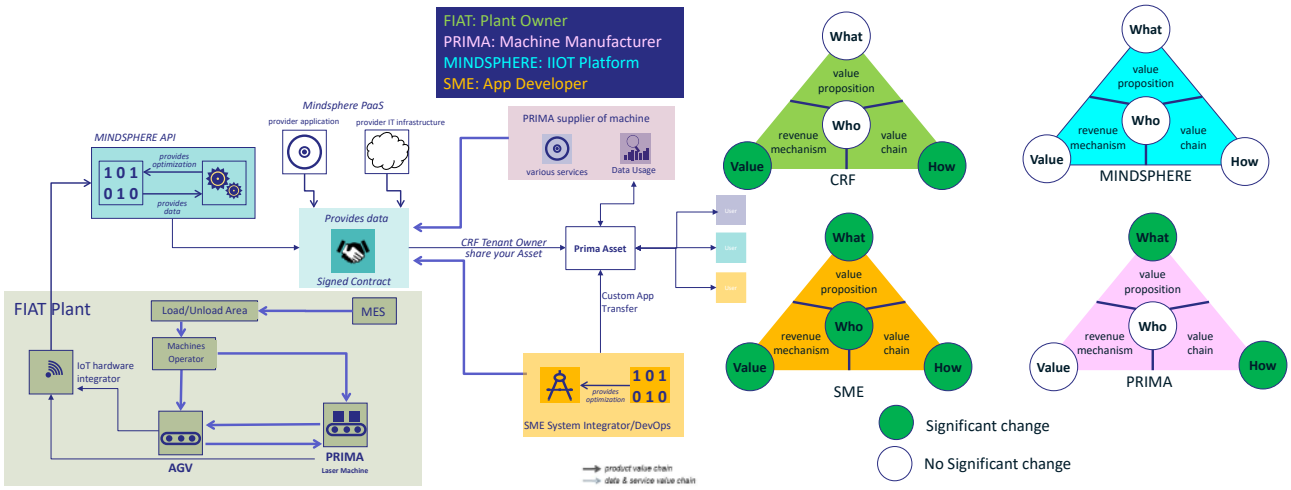


Figure 17 - Boost4.0 FIAT use case of Data Exchange in a close platform

Another more classical example of Data Spaces in Manufacturing is the Boost4.0 Lighthouse pilot in **Whirlpool** EMEA. Here, in a more traditional way, the point is to optimise the stock levels of the washing machines spare parts, by analysing a huge and heterogeneous amount of data coming from different Whirlpool departments. Here the point is how to distribute along the value chain (up to the suppliers of spare parts and down till to the maintenance workers with their vans visiting homes) the results of an AI-based forecast by disclosing just the necessary information with the due explanations and not the most confidential information at the basis of the forecast.

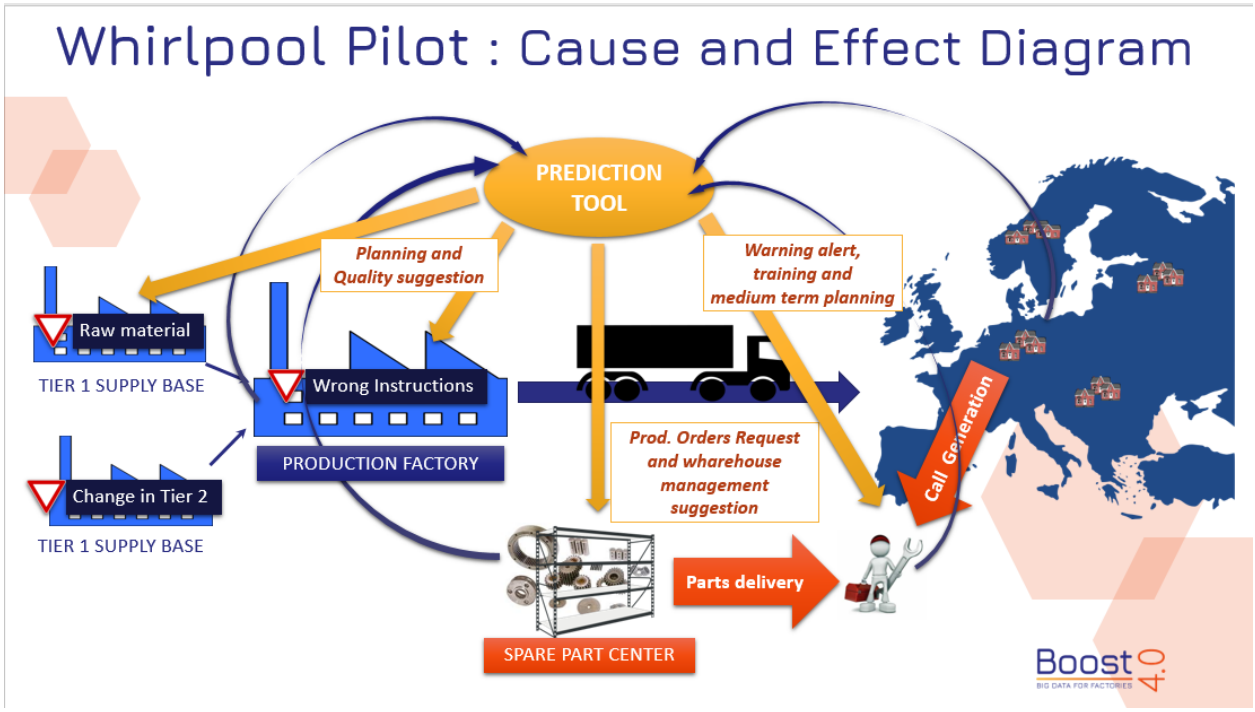


Figure 18 - Boost4.0 WHIRPOOL use case of Data Exchange in an open platform

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

4 Materialising Data Spaces in the Manufacturing Sector

While in Chapter 3 we have analysed the Business side of the "**Data Spaces for Manufacturing**" coin, in this chapter 4 we are going to describe some approaches to the technical implementation of it. The **European Commission** recently held some workshops on the matter along 2020. More precisely on November 23rd 2020 a webinar[20] (§4.1) focussed on the state of play and anticipated the DEP call about Data Spaces. **BDVA** is approaching the topic from a cross-domain perspective and its two editions of the report "Towards a common European Data Sharing Space" were released in 2019[21] and 2020[22] (§4.2). **IDSA** is of course leading the implementation side of the challenge and several working groups have been activated for the definition and materialisation of the so-called Building Blocks (§4.3). Thanks to the **OPEN DEI CSA[23],** a cross-domain and cross-disciplinary Task Force has been created to identify and study the design principles of Data Spaces and a paragraph of the conclusive report (under preparation at today) is devoted to the Manufacturing sector (§4.4). In particular, the Manufacturing community behind OPEN DEI has elaborated a **Correspondence Matrix** where to cross the CF1 pathways (smart autonomous factory, collaborative product-service factory, hyper-connected factory) with the business models described in §3.2 (open data, data marketplace, trusted data network). This is reported in §4.5.

4.1 The European Commission perspective

In the recent EBDVF2020 event (23rd November 2020) the state of play about Manufacturing Data Spaces was positioned by the European Commission. As affirmed by the EU Data Strategy: "The Commission will promote the development of common European data spaces in strategic economic sectors and domains of public interest. This should lead to the availability of large pools of data in these sectors and domains, combined with the technical tools and infrastructures necessary to use and exchange data, as well as appropriate governance mechanisms."

Focus on sharing, pooling, and reusing data across organisations in the manufacturing sector, the objective in the EC perspective is to set up and deploy several operational data spaces (**Embryonic Data Spaces**) for specific value chains in the manufacturing sector, which enables companies in different user roles (supplier, client, service provider, …) to interact with large amounts of manufacturing data.

The Commission-funded study (2018)[24] on data sharing by companies in Europe analysed in detail the EU landscape and specifically focusses on the barriers and challenges for Data Share and Data Re-use.

Regarding Data Sharing, out of a sample of 129 EU companies, the study revealed that 39% of them do share data, but more than a half of them are facing obstacles to share. In particular, Technical (73%), Legal (54%), Control (42%) and Skills (38%) obstacles have been identified.

[20] https://ec.europa.eu/digital-single-market/en/news/data-spaces-manufacturing-current-state-play
[21] Microsoft Word - BDVA DataSharingSpace PositionPaper_April2019_V1.docx
[22] Microsoft Word - BDVA DataSharingSpaces PositionPaper V2_2020_Final.docx
[23] https://www.opendei.eu/
[24] https://ec.europa.eu/digital-single-market/en/news/findings-commission-funded-study-data-sharing-companies-europe

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

Regrading Data Reuse, in the same sample above, the percentages are quite similar with 42% that reuse data and 59% which face obstacles. In this case, the most voted obstacles have been: Access denial (66%), Unfair Conditions and High Costs (41%) and lack of interoperability (31%).

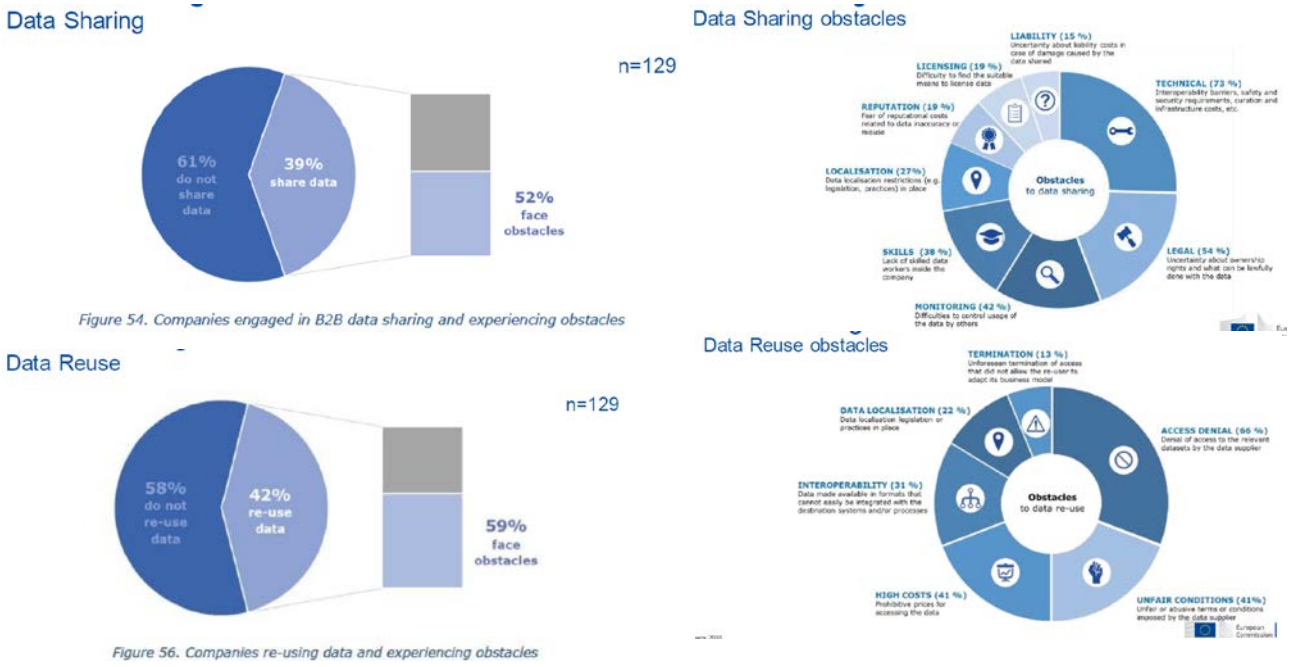


Figure 19 - EU study about barriers and challenges for Data Share and Data Re-use, November 2020

In the subsequent (Feb 2020) EU Data Strategy, more or less the same or similar problems have been detected, such as **Not enough data available for reuse** (more public sector data, low uptake of voluntary data sharing in companies, no clarity in the use of private sector data for the common good, This has been addressed by the Data Governance Act reported above), **Lack of a European data processing and storage solutions** (to inspire a pan-EU solution such as Gaia-X, see below), **Absence of comprehensive data governance approaches** (standardisation and interoperability across sectors), **Skills shortage and low data literacy** (which inspired our CF2 Data Space pathway in Chapter 5), **Fragmentation of the Single Market** and **No real user empowerment**.

During the SMI workshop at the recent **European Big Data Value Forum** (Manufacturing Track, November 4th 2020)[25], the Commission identified the main technical (in yellow) and main applicative building blocks (in blue for the AI domain) materialising the Data Space for Manufacturing concept. Among the Yellow technical boxes, we can mention High Value Data Sets, Data and Cloud Infrastructures and Services and a European AI on-demand platform, while among the Blue ones, we mention a Manufacturing-driven Data Space, the AI Testing and Experimental Facilities for Manufacturing and several manufacturing advanced applications. The availability of data from individual companies is of course a precondition.

[25] https://www.european-big-data-value-forum.eu/manufacturing/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

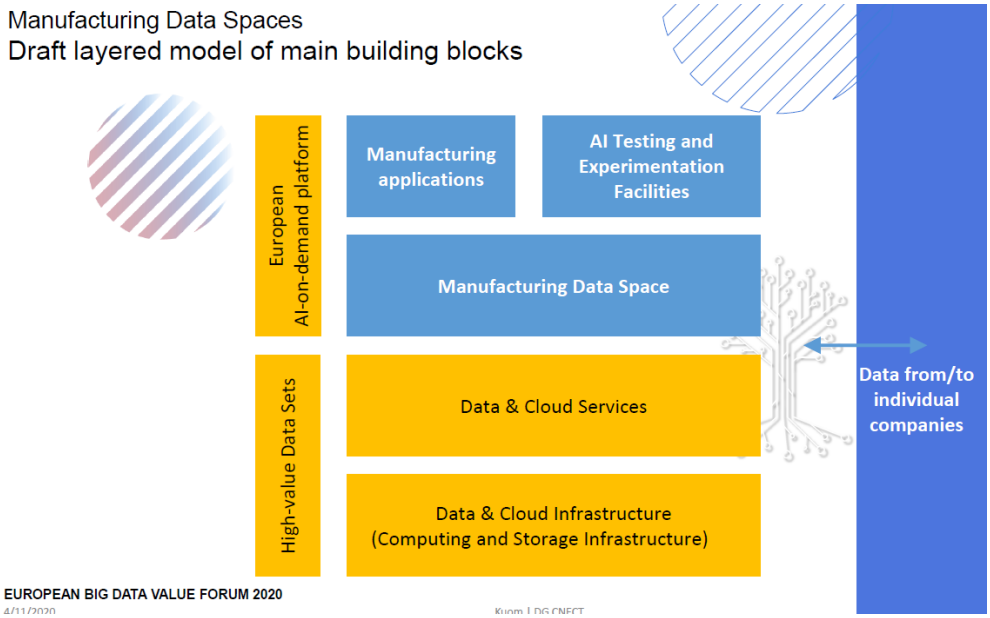


Figure 20 - European Big Data Value Forum: Model for manufacturing Data Spaces

Moreover, in the November 23rd workshop, the Manufacturing Data Spaces as to be implemented in the DEP workprogramme (2021-2022) has been presented. Two verticals have been selected for projects: agile supply chain management and execution (in particular, Manufacturing as a Service and one-of-a-kind production) and dynamic assets management and Predictive Maintenance.

These two topics are quite related to the Hyper-connected and the Smart Autonomous Factory pathways of CF1. A more horizontal topic has been also presented and is concerned with the Green Deal and the need for product and material data sharing along the lifetime of complex products, in order to support circularity, recycling and de- re-manufacturing).



Figure 21 - Manufacturing Data Spaces, DEP workprogramme (2021-2022)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

4.2 The BDVA Perspective: Towards a common European Data Sharing Space

Inside BDVA, in April 2019 a new Position Paper has been released "Towards a European Data Sharing Space – Enabling data exchange and unlocking AI potential" where more clear definitions about Data Sharing Spaces have been given as well as some future recommendations on how to implement it.

In particular, the following definitions were given:

Data Spaces: any ecosystem of data models, datasets, ontologies, data sharing contracts and specialised management services, together with soft competencies around it, enabled by a **Data engineering** approach.

Data Platforms: architectures and repositories of interoperable hardware/software components, which follow a software engineering approach to enable the creation, transformation, evolution, curation and exploitation of both static and dynamic data in Data Spaces.

Data Marketplaces: they offer a one-stop-shop where data providers (sellers) and data consumers (buyers) can meet, match and trade their respective data assets and requirements, as well as data algorithms and processing approaches. Thus, Data Marketplaces operate on top of Data Spaces and utilize Data Sharing and Data Exchange apps provided by Data Platforms. At the basis of Data Marketplaces there is the concept of valorisation and monetisation …*for successful Development, Implementation, Adoption of a European Data Sharing Space that: allows new and existing vertical, cross-sectoral, personal and industrial data spaces to interoperate; offers services and experimentation opportunities to all stakeholders; promotes European values…*

Among the recommendations, we can mention:

1. Create Conditions for development of a trusted European data sharing framework. This is mostly directed to **EU policy makers** (in order to align local/national strategies to realise framework for exploitation by industry) and to **Member States** (in order to showcase evidence-based benefits for diverse EU stakeholders). Critical matters identified are: Data rights management/Sovereignty/Usability/Governance;

2. Incorporate data sharing at core of data lifecycle (for greater access to data), which implies the development of Data life-cycle management strategies that natively incorporate data sharing events and transactions and could include interoperability, security, quality, privacy and ethics by-design;

3. Support EU businesses to safely embrace new technologies, practices and policies by developing Facilities and environments offering safe testing and de-risking of new business and innovation models and by exploiting existing mechanisms (DIHs, i-Spaces) as well as new instruments (e.g., European-wide Regulatory Sandboxes)

4. Assemble EU-wide digital skills strategy to equip workforce for new data economy, in order to revisit educational policies holistically to reskill workforce and prepare professionals for dynamic careers and to develop strategies to re-define or develop new jobs and abilities (e.g., judging bias, assessing liability).

The 2020 edition of the same Position Paper is instead focussing on a new reference model for Data Sharing Spaces: the **Data Sharing Value Wheel**. As depicted in the figure below, the success of widespread data sharing activities revolves around the central key concept of trust: in the validity of the data itself and the algorithms operating on it; in the entities governing the data space; its enabling technologies; as well as in and amongst its wide variety of users (organisations and private individuals as data producers, consumers or

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

intermediaries). To achieve the required levels of trust, each of the following five pillars must meet some of the necessary conditions.

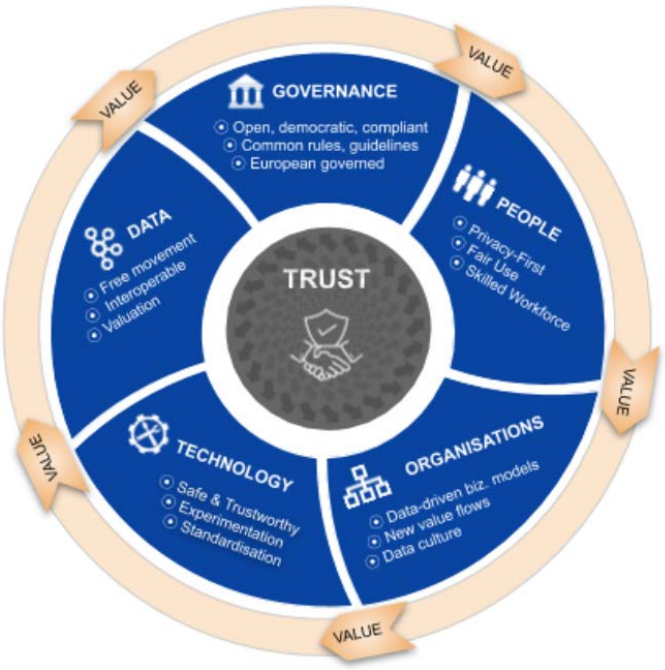


Figure 22 - The European Data Sharing Value Wheel

1. **Data** - As a touted 5th European fundamental freedom, free movement of data relies on organisational data strategies that embed methodologies for data sharing by-design (e.g., interoperability) and clear standard guidelines that help determine market value of data assets.
2. **Governance** - A European-governed data sharing space can inspire trust by adhering to the more advanced European rules, guidelines and regulations and promote European values. Participation should be equally open to all and subject to clear and fair rules of conduct.
3. **People** -Data sharing needs to guarantee individual privacy and offer fair value or compensation of shared personal data. In order for Europe to drive data sharing activities, the European workforce needs appropriate reskilling and upskilling to meet the evolving labour market's needs.
4. **Organisations** - More organisations (including business, research and governmental) need to rethink their strategy to fully embrace a data culture that places data at the centre of their value proposition, exploring new data-driven business models and exploiting new data value flows.
5. **Technology** - Safer experimentation environments are needed to catalyse the maturation of relevant technology behind trustworthy data, data access and algorithms (privacy, interoperability, security, and quality). Standardisation activities need to adjust for faster reaction times to emerging standards and the identification of new ones.

Twelve new recommendations have been also identified and are summarized here below:

1. **Launch Coordination actions** to map existing initiatives at EU, member state, regional and municipal level (leveraging existing EC CSAs), and identify the common building blocks to create an impact and foster standardisation actions (see action no. 6). Efforts should particularly seek to collect, publish and

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

systematically analyse use-cases (see action no. 8) and align legal (including tax treatments) and governance structures necessary for safe and fair level European-wide data sharing (see action no. 5).

2. **Further invest in a Euro-wide Skills Strategy** to identify major skill and expertise gaps in the European workforce, devise Euro-wide reskilling and upskilling roadmaps, advise education and training bodies with remedial actions and offer further scholarships and fellowships in crucial areas of study.
3. **Funding the development** of rules and guidelines for iterative, multi-phase design, creation, scaleout or merging **of existing data sharing spaces**.
4. **Launching Research and Innovation Actions** (including EC R/IAs) to evolve high-impact technology (privacy, interoperability-, security-, quality- and ethics-by-design) with an emphasis on an interoperability initiative across existing spaces and the deployment of trust solutions for data sharing assurance.
5. **Developing EU-wide Data Governance practices** to clearly define rules of conduct for the fair use of exchanged data after testing different approaches to assess the impact of data sovereignty, campaign for the revision of relevant legislation, and explore means for Conformity Assessment (e.g. voluntary or licensed certification, data sharing assurance agreements) that guarantee trust.
6. **Proactive coordinated engagement with international standardisation bodies** to establish Key Performance Indicators (KPIs) that drive standardisation of common building blocks (see action no.1), successful technology (see action no.4), methods (e.g., data rights management as ownership solution) and best practices (e.g., guidelines for international data valuation, privacy preservation and FAIR principles for non-research data) behind trust-enabling solutions.
7. **Assurance of sufficient conditions** (assurance, market demand, government procurement) to enable widespread adherence to established standards, rules and guidelines and promotion amongst networks.
8. **Investing in piloting** to test-drive identified data sharing use-cases in safe and dynamic regional and European-wide cross-sectoral scenarios (pairing member states with different levels of progress in data sharing infrastructures).
9. **Engage with major initiatives offering experimentation** activities that rely on data sharing (EDIH future network, BDVA i-Spaces network, etc.) to jointly explore market capabilities for sustainable new business, cross-sectoral opportunities and organizational models (e.g., for data valuation, and/or organizational transformation) and create innovative solutions.
10. **Set up European regulatory sandboxes** for trusted data sharing and safe exploration of risks, so as to encourage innovation and motivate data sharing.
11. **Launch a campaign targeted at influencing organisations** to revisit their data strategy and lifecycles to ensure the production of data that is sharing-ready by design.
12. **Launch a EU-wide citizen-oriented campaign** for an open, democratic and fair data economy and the right to the free flow of data that is both safe and trusted. Drivers of activities for greater awareness include the EC, government at member state, regional and municipal level, sector-based industrial associations, entrepreneurs and technology providers and consumer/end-user organisations.

4.3 The International Data Spaces Association perspective

The perspective of IDSA is summarised in series of images taken from recent presentations of IDSA representatives. First of all, the notion of a European Data Space is divided into three main blocks: a) availability of high Value pools of data sets (following open standard data models and ontologies); b) availability of an infrastructure to use and exchange data (industrial data platform in our case) and c) availability of appropriate governance mechanisms.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

This will allow the domain specific implementations for the nine critical sectors of Manufacturing, Health, Agriculture, Green, Finance, Public, Mobility, Energy and Skills.

In a layered architectural model (see picture below), the Data Sharing Infrastructure (point b) above) is located between the Cloud/Edge infrastructure and the ecosystem of services. Such infrastructure is able to provide Essential Trust Services (e.g. certification and clearing house) and Basic Data Services (app store, broker, data connectors).

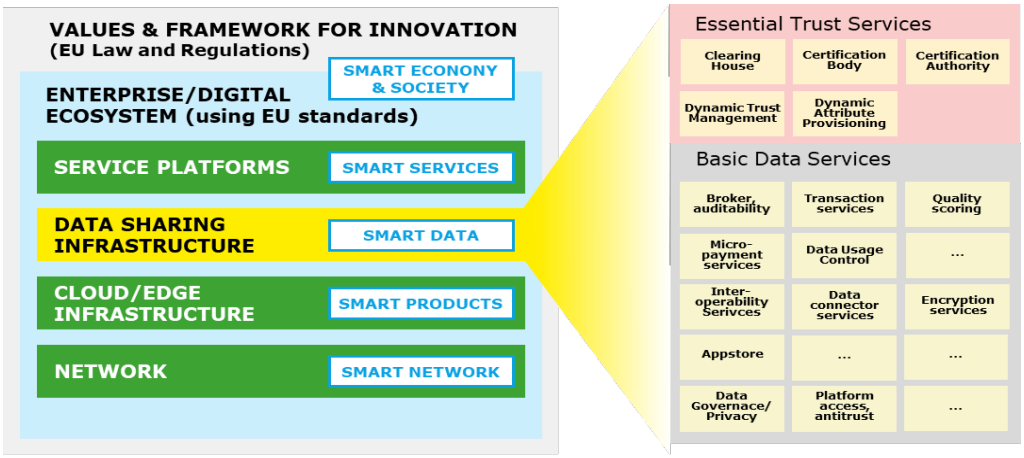


Figure 23 - IDSA layered architectural model

Another complementary view of the Data space is in a collection of generic and standard models to support Endless Connectivity (Data Standards, Metadata and Data Exchange tools), Trust between different security domains (AAA and certification), Governance form the data economy (agreements, earning models, certification and monitoring

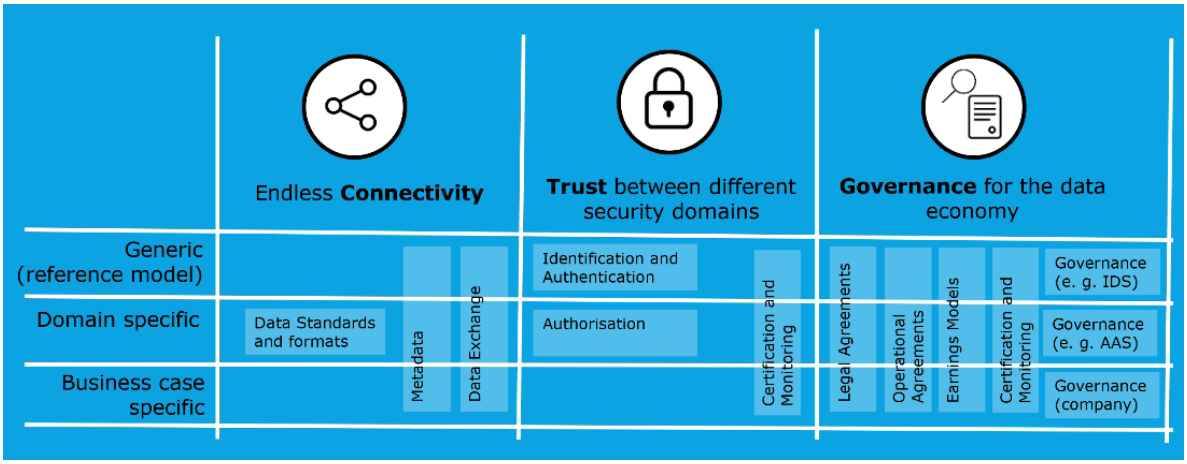


Figure 24 - Data Space elements from the point of view of Connectivity, Trust and Governance

The reference implementation model indicated in the IDS reference architectural model is still valid for any application domain and sector.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

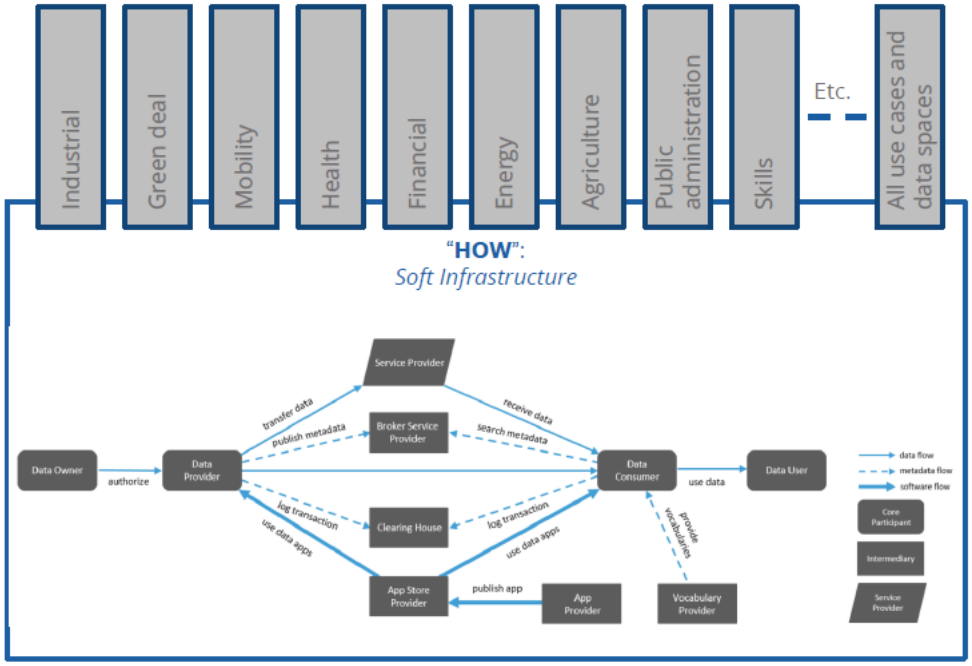


Figure 25 - IDS reference implementation model

Lastly, a key point ion the IDSA vision is the X convergence between the Data and the Infrastructure ecosystem that form together the GAIA-X schema,

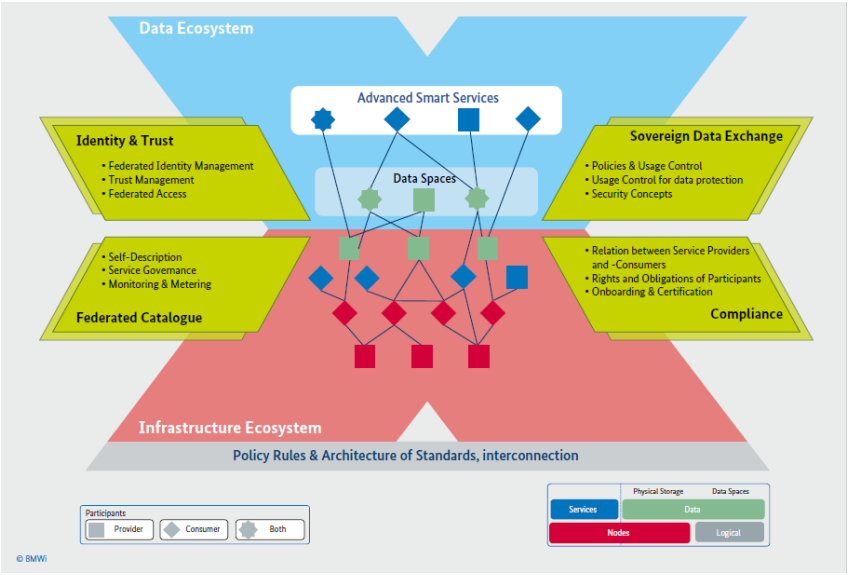


Figure 26 - High level overview of the GAIA-X architecture

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

4.4 The OPEN DEI perspective: Task Force 1 Design Principles for Data Spaces

OPEN DEI Task Force 1 (TF1) "Design Principles for Data Sharing Spaces" is currently finalising its position paper. The OPEN DEI Digital Transformation Data Sharing Spaces task force has the aim to develop a common understanding, approach and design principles among the four OPEN DEI domains (manufacturing, energy, agrifood and health & care) towards the development of a common European Data Space[26].

In the above domain, one of the most advanced, concrete, mature initiatives in Europe is IDSA[27], who is going to lead this TF, but of course we welcome in the discussion also other voices and opinions especially from the four domains. In this perspective, we share in the TF1 repository facilitated by IDC, a collection of basic IDSA documents including the IDSA RA 3.0, the standard for Secure Gateways for Data and Services DIN SPEC 27070 and the most recent presentations and papers such as Lars Nagel (IDSA CEO) presentation during OPEN DEI webinar on April 30th 2020 and IDSA's position paper "Implementing the European Strategy on Data".

In particular, the EU Data Strategy, published on February 19th 2020, identifies four main pillars to be addressed in order to get to a common EU Data Space[28]: a legal framework for data access and use, a technical infrastructure and building blocks, a capacity building program for skills and competencies, a rollout in crucial economic sectors and domains of public interest. The following picture illustrates this last pillar and the need for a common approach and technical baseline among the four selected OPEN DEI domains.

In the context of OPEN DEI Task Force 1, IDS Association provides: Open **Reference architecture** setting the framework conditions – based on standards; **Trust framework** and scheme for data sharing; Coordination operations for essential **data sovereignty services.**

The OPEN DEI task force 1 is discussing how to implement this data spaces in four of the above domains: manufacturing, health & care, agri-food and energy. One of the key challenges of the TF1 is to discuss and agree on a set of coherent and coordinated design principles for data spaces design, development and rollout, able to support and guarantee data sovereignty, i.e. "a natural person's or corporate entity's capability of being entirely self-determined with regard to its data[29]". We have identified three main Building Blocks (Connectivity, Trust, Governance) which will drive us in our endeavour. In each of them, we have mapped the nine Building Blocks for Data Sharing Spaces identified in a recent study by InnoPay[30].

Objectives:

- Analyze and discuss **requirements and common challenges** for data sharing spaces
- Coordinate and work together on needs for **cross-sectorial governance for data access and use** based on common principles agreement and market-ready enabling building blocks.
- **Share knowledge, best practices** and create a network of cross-sectorial data sharing stakeholders.

[26] https://ec.europa.eu/digital-single-market/en/news/communication-towards-common-european-data-space

[27] https://www.internationaldataspaces.org/

[28] I want European businesses and our many SMEs to access high quality data and create value for Europeans including by developing AI applications. Thierry Breton, Commissioner for the Internal Market

[29] https://www.isst.fraunhofer.de/content/dam/isst-neu/documents/Publikationen/StudienundWhitePaper/FhG-ISST_DATA-ECOSYSTEMS.pdf

[30] https://www.innopay.com/en/publications/how-are-soft-infrastructures-ishare-promoting-data-sovereignty-create-new-value

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

- Create **awareness** of task force and IDSA activities:
  - Challenges and awards
  - Dissemination activities
  - Joint paper
- Develop **recommendations for upcoming policies** of EC and MS to support data exchange initiatives and for standardisation activities.

Task force management:

IDSA will coordinate the TF1 with the support of Technical coordinator and OPEN DEI WP2-WP3 and WP5 leaders, ENGINEERING, ATOS and FIWARE Foundation. The ambassadors will propose and establish the first contact with representatives of its sector. Furthermore, there will be a coordination with the working groups in each sector where building blocks identification and mapping activities are being developed. There will be a directory on OPEN DEI Box for sharing documents and a distribution list and a Teams channels that facilitates the communication process too.

Each meeting we will produce a resume of the attendees, main conclusions. After each meeting we will distribute a summary and agreements. IDSA will evaluate the conclusions and contributions from the task force to make recommendations to EC.

Task force (13) experts from each sector:

- Agrifood sector:
  - **DEMETER:** Srdjan Krco from DunavNet
  - **ATLAS:** Mallku Caballero from AgriCircle
  - **IoF2020:** Harald Sundmaeker from ATB
- Health care sector:
  - **Smart4Health:** Ricardo Goncalves and Maria Marques from Uninova
  - **InteropEHRate:** Francesco Torelli and Alessio Graziani from Engineering
  - **PHArA-ON :** Andre Grguric from Ericsson.
- Energy sector:
  - **BD4OPEM**: Javier Valino -from ATOS.
  - **InterConnect**: Antonio Kung from Trialog
  - **PLATOON** Eugenio Perea from Tacnalia
- Manufacturing sector:
  - **KYKLOS**: Kjell Bengtsson from Jotne.
  - **DigiPRIME**: Luis Usatorre from Tecnalia
  - **DigiPRIME**: John Soldatos from Intrasoft Int.
  - **Qu4lity**: Jorge Rodriguez from Atos

Task force outcomes:

The iteration 1 from September to December 2020 has the target to achieve the following results:

- Summary of each task force meeting and two task force activities reports for the EC.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

- An OPEN DEI task force **OPEN DEI cross-sectorial position paper** about design principles and components for common EU data spaces and their maturity level with the active participation of selected representatives from the OPEN DEI ecosystem and the four domains Innovation actions.
- Joint presentation of cross sectorial position paper and sectorial dissemination activities plan.
- IDSA will inform EC about the task force activities. Furthermore, the relevant conclusions will be reported to the EC.

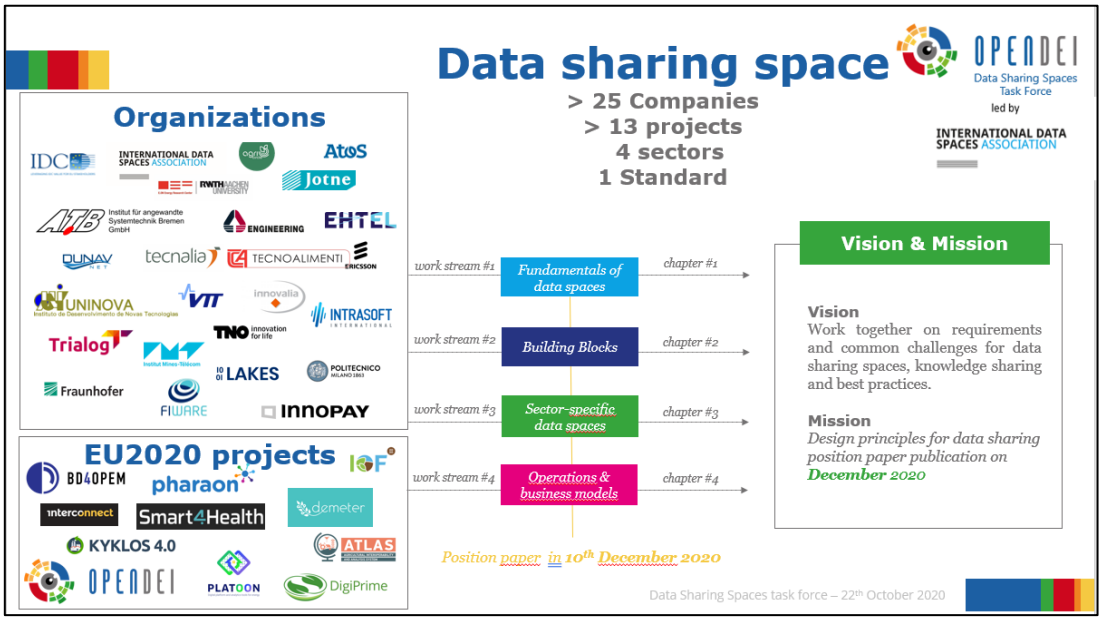


Figure 27 - OPEN DEI study on European Data Sharing Space

The TF1 paper is organised in four main central chapters: Fundamentals of Data Spaces, Building Blocks, Sector-specific Data Spaces and Operations and Business Models and sees the collaboration of 43 experts from the four domains.

Regarding Data Spaces for Manufacturing in Chapter 3, the paper aims at achieving four main objectives: From mass production to mass customization; Reliable supply chains; Servitization and Circular manufacturing. All these objectives rely on data sharing and cannot be accomplished without the manufacturing sector digitalization. Several efforts have been done in order to define the digitalization pathway for different types of factories. On the lower level is the data capture while on the upper level is the data analysis and decision support systems. RAMI-4.0 (Reference Architecture Model for Industrie 4.0) has analyzed the data transmission both horizontally and vertically.

The existing situation in the manufacturing sector is that, on the one hand, information -and data- remain in silos in the different factory departments; and on the other hand, there is no available open data related with the manufacturing processes. Manufacturing companies are nowadays analysing the benefits of data spaces and the data economy. Also, the Data Economy offers new Business Models that are challenging todays way of doing business so Manufacturing companies need to explore new approaches. It has also to be taken into account that the more manufacturing companies go digital, the higher they are exposed to cyber-criminality.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

For the time being, data analytics is being used mainly in the following fields: Image processing and computer vision; Predictive maintenance and failure detection; Quality control and equipment diagnosis; No linear control and robotics. Manufacturing companies consider the impact of data in the following aspects: Cost reduction, Quality improvement, Servitization and Supply chain management.

As one example of front-runners, the **Smart Connected Supplier Network** (SCSN) is an initiative of manufacturing companies and their IT-suppliers in the high-tech manufacturing supply chain. The aim is to facilitate cross-factory communication and thus ensuring supply chain transparency and interoperability. Manufacturing companies can use a single SCSN connection to exchange purchase-to-pay information with all their suppliers and customers. This reduces administrative efforts and human errors while increasing the supply chain agility.

4.5 A Reference Matrix for Data Spaces in Manufacturing

In synthesis we can position **Manufacturing Data Spaces** (MDS) in the following reference matrix, where existing and future cases can be mapped in. Manufacturing Data Spaces could therefore be developed in **Autonomous Smart Factories** with the aim to collect and exchange data from autonomous AI-based Manufacturing Assets in the shopfloor, for instance by exploiting the RAMI4.0 Asset Administration Shell[31] standard; in **Hyper-connected Factories** with the aim to digitalise very dynamic value chain ecosystems and support Manufacturing as a Service and Additive Manufacturing fablabs; in **Collaborative Product Service Factories** to follow all the phases of the lifecycle of a complex Product Service System till its end-of-life in a Circular Economy perspective, through the interaction between Digital Twins of Things and Digital Twins of Humans (Personas).



Figure 28 - Reference matrix for Data Space in Manufacturing

[31] https://www.plattform-i40.de/PI40/Redaktion/EN/Downloads/Publikation/vws-in-detail-presentation.pdf?__blob=publicationFile&v=11

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

5 The CF2 Data Spaces pathway

5.1 Data Spaces levels and Dimensions

But, what do we mean for a common EU Data Space? CF2 pathway invites to refer to BDVA publications and in particular the April 2019 (under revision currently) position paper "**TOWARDS A EUROPEAN DATA SHARING SPACE** Enabling data exchange and unlocking AI potential"[32] where Data Spaces, Data Platforms and Data Marketplaces are defined inside the overall EU Data Ecosystem and to the BDVA "**SRIA Strategic Research and Innovation Agenda** version 4.0"[33] reference model where five plus one data-oriented challenges have been identified and materialized.

g. ***Data Management**: Principles and techniques for data lifecycle management including quality, integrity, provenance and liability*
h. ***Data Protection**: Privacy and anonymisation mechanisms to facilitate data protection. This is shown related to data management and processing as there is a strong link here, but it can also be associated with the area of cybersecurity.*
i. ***Data Processing Architectures**: Optimised and scalable architectures for analytics of both data-at-rest and data-in-motion, with low latency delivering real-time analytics*
j. ***Data Analytics**: Data analytics to improve data understanding, deep learning and meaningfulness of data, supporting forecasting, prediction, prescription and autonomy*
k. ***Data Visualisation and User Interaction**: Advanced visualisation approaches for improved user experience*
l. ***Data Sharing along Value Ecosystems** and Innovation Support: Data platforms for data sharing include, in particular, Industrial and Personal Data Platforms, but also other data sharing platforms like Research Data Platforms (RDPs) and Urban/City Data Platforms (UDPs).*

These six dimensions have been chosen for the analysis of "Data Space" maturity of a Manufacturing company (SME), along five levels of maturity defined as:

vi. **No Data Control**. This is a level where companies do not have CONTROL of DATA along their lifecycle. DATA are of course generated but no models, no processes, no awareness is in place. Smart Machines, Smart Products and Smart Value Chains do produce data but they are dispersed in several areas and the company does not have any control on them.
vii. **Data Silos.** This is a level where the company understands the relevance of Data lifecycle management as for feeding Enterprise applications such as PLM ERP SCM CRM MES. Data are captured, protected, processed, analysed, visualized in the context of their Enterprise application, so some data is NOT managed, some data is duplicated, but in any case DATA is SILOED.
viii. **Data Bridges.** This is a level where for feeding some advanced applications in the domain of planning, optimisation, prediction, the company needs to integrate data from different sources (e.g. ERP-CRM; ERP-MES, PLM-MES). Ad-hoc integration bridges are developed but without an holistic and comprehensive strategy for data management and governance.

[32] https://www.bdva.eu/sites/default/files/BDVA%20DataSharingSpace%20PositionPaper_April2019_V1.pdf
[33] https://www.bdva.eu/node/874

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

ix. **Data Interoperability.** This is a level where enterprises define a strategy and a set of processes for managing and governing the Data generated inside and acquired from outside along their complete lifecycle. Data Spaces are generated where common standard data models, ontologies, industrial data platforms are set up inside the company and along the product lifecycle and the value chain. Any new application can be easily plugged & played in this data space just by configuration, so that scalability is assured.

x. **Data Valorisation.** This level is when the enterprise fully understands how data can be profitably exploited inside and outside of the value chain. Data Marketplaces and/or Trusted Data Networks are setup and managed in agreement with Data Protection (personal data) and Data Sovereignty (non-personal data) principles. This does not necessarily imply a monetisation and revenue stream, but the full exploitation of the intrinsic value of data-information-knowledge.

Now the DS pathway 6*5 contribution space is being filled by contributions of projects and ZDM projects are required to provide its view in the domain of Data Spaces for Zero Defect Manufacturing in Autonomous Quality scenarios. ZDM industrial pilots will be soon required to compile an online survey and to be interviewed about their maturity regarding Data Spaces. Such a maturity analysis could also be interpreted as an action to improve Data Culture and Data Value Awareness also in the manufacturing industry, where Data Economy principles are now leading new business and innovation models in other domains such as Mobile Communication, Retail and Finance.

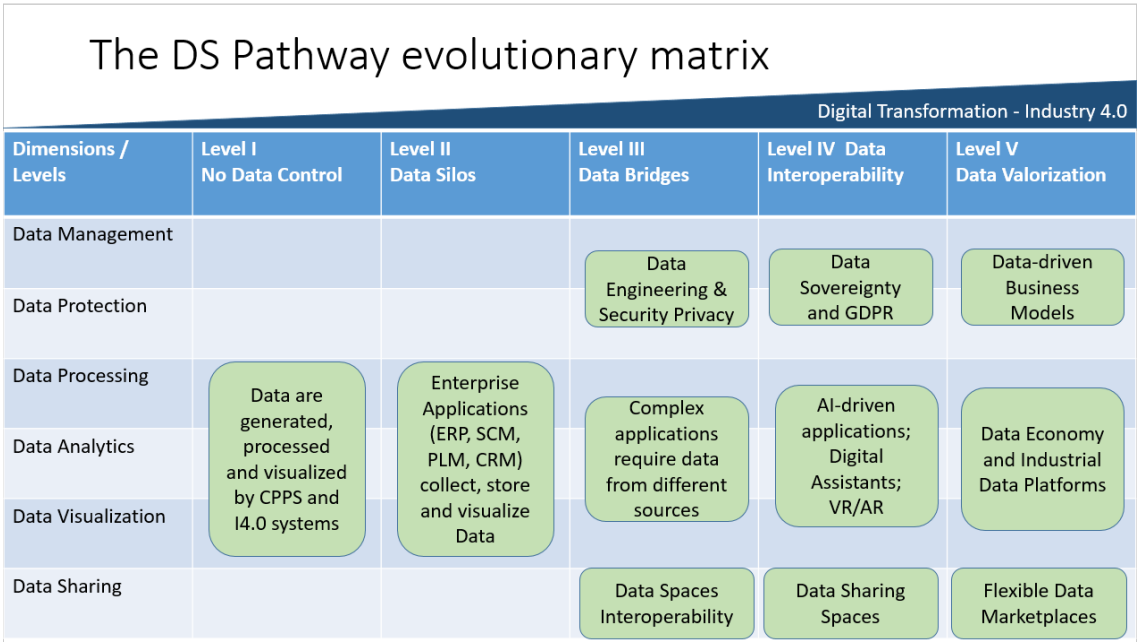


Figure 29 - Data Space Evolutionary Matrix

5.2 The Data Spaces Pathway Survey

As already mentioned, the starting point to evaluate the position on the Data Space Pathway is a self-assessment of the current situation, taking into consideration all the different areas related to data. To drive the enterprise in this task, it has been realized an online survey in order to provide an overview of the

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

company maturity regarding Data Spaces. It will be followed by a face-to-face interview, that will allow to deepen and explore the same aspects but with a more detailed and precise approach.

The survey, available at this link, is based on the evolutionary 6*5 matrix described above: it consists of 6 questions (one for each dimension), admitting 5 possible answers (one for each level). Namely, questions are:

i. ***Data Management**: To what extent is your company implementing principles and techniques for data lifecycle management including quality and integrity?*

- *LEVEL_a.* ***NO DATA CONTROL** There is no Data Management culture in the company*
- *LEVEL_b.* ***DATA SILOS** Data are stored in Enterprise Applications and managed from within such systems*
- *LEVEL_c.* ***DATA BRIDGES** Cross-application Data Management Bridges are created on an ad-hoc basis*
- *LEVEL_d.* ***DATA INTEROPERABILITY** A common set of Data Management applications is applied systematically to all the Data Sources and Enterprise Applications*
- *LEVEL_e.* ***DATA VALORISATION** Data Management applications (Data Quality in particular) are used to assign a value to data along its life cycle*

ii. ***Data Protection**: To what extent is your company implementing privacy and anonymization mechanisms to facilitate data protection (cybersecurity)?*

- *LEVEL_a.* ***NO DATA CONTROL** No specific Data Protection measures in place regarding storage, computation and transfer of data*
- *LEVEL_b.* ***DATA SILOS** Each single Enterprise Application (ERP SCM PLM) is responsible for Data Protection measures and actions*
- *LEVEL_c.* ***DATA BRIDGES** Secure Data Bridges between silos are implemented on a case-by-case basis*
- *LEVEL_d.* ***DATA INTEROPERABILITY** Common Data Protection mechanisms and tools are shared at company level and used in every application*
- *LEVEL_e.* ***DATA VALORISATION** Data Protection mechanisms associated to Data Spaces increase the value of Data*

iii. ***Data Processing Architectures**: Is your company implementing performant and scalable data-driven architectures for analytics of both data-at-rest and data-in-motion?*

- *LEVEL_a.* ***NO DATA CONTROL** Data is not processed by the company or it is processed by default procedures (of the suppliers) out of control for the company*
- *LEVEL_b.* ***DATA SILOS** Data is processed inside Enterprise Applications (ERP SCM PLM) according to architectures and tools predefined by them*
- *LEVEL_c.* ***DATA BRIDGES** Data Processing architectures (embedded, edge, cloud) are implemented for ad-hoc data bridges in Enterprise Applications*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

LEVEL_d. ***DATA INTEROPERABILITY*** *A common Data Processing architecture is adopted for all Enterprise Applications including both Data at Rest and Data in Motion processing*

LEVEL_e. ***DATA VALORISATION*** *Data Processing tools are associated as a service to Data Sets, increasing their value in the Data Space and implementing Data Economy services*

iv. ***Data Analytics****: To what extent is your company implementing data analytics to improve data understanding, deep learning and the meaningfulness of data?*

LEVEL_a. ***NO DATA CONTROL*** *Data analytics not performed or performed by third parties out of control of the company*

LEVEL_b. ***DATA SILOS*** *Data Analytics procedures are implemented (and replicated) from inside Enterprise Applications only*

LEVEL_c. ***DATA BRIDGES*** *Data Analytics procedures are implemented across silos in an ad-hoc fashion*

LEVEL_d. ***DATA INTEROPERABILITY*** *A common agreed repository of Data Analytics tools is shared in the company, at disposal for any application*

LEVEL_e. ***DATA VALORISATION*** *Data Analytics tools are associated as a service to available Data Sets, increasing their value in the Data Space and implementing Data Economy services*

v. ***Data Visualisation and User Interaction****: How is your company adopting advanced visualisation approaches for improved user experience including VR/AR?*

LEVEL_a. ***NO DATA CONTROL*** *Data visualisation tools not implemented or implemented by third parties out of control of the company*

LEVEL_b. ***DATA SILOS*** *Data visualisation tools embedded in Enterprise Applications on top of predefined sets of Data*

LEVEL_c. ***DATA BRIDGES*** *Data visualisation applications are developed ad-hoc cross-silos to show data coming from different sources*

LEVEL_d. ***DATA INTEROPERABILITY*** *A common agreed repository of Data Visualisation tools is shared in the company, at disposal for any application*

LEVEL_e. ***DATA VALORISATION*** *Data Visualisation tools are associated as a service to available Data Sets, increasing their value in the Data Space and implementing Data Economy services*

vi. ***Ecosystems for Data Sharing*** *To what extent is your company implementing secure and trusted Data Sharing spaces and platforms along the value chain, according to Data Sovereignty principles?*

LEVEL_a. ***NO DATA CONTROL*** *Data Sharing is not implemented or implemented without any control*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

LEVEL_b. ***DATA SILOS*** *Data Sharing is allowed just from inside Enterprise Applications to a predefined value ecosystem*

LEVEL_c. ***DATA BRIDGES*** *Data Sharing from different sources are ad-hoc implemented for specific suppliers and customers*

LEVEL_d. ***DATA INTEROPERABILITY*** *A common and agreed architecture and set of tools for secure and sovereign Data Sharing is implemented in the company*

LEVEL_e. ***DATA VALORISATION*** *Data Sharing applications are associated as a service to available Data Sets, increasing their value in the Data Space and implementing Data Economy services*

5.3 The Data Spaces Pathway Interactive Board (MURAL)

Besides the online survey, whose main goal is to evaluate where the enterprise is positioned in the Data Space Pathway, we have designed an interactive board based on the evolutionary matrix displayed in Figure 1.

The objective is to make participants reflect about following topics, related to Data Spaces:

- **Obstacles and challenges** → Which are the barriers that keep manufacturing enterprises stuck to the lowest levels? Or more in general, which obstacles do enterprises face when they want to improve their "data maturity"?
- **Industrial cases** → According to personal experience, which projects/activities (and in which way) have supported and stimulated the implementation of Data Space in Manufacturing?
- **Future Opportunities** → Which are the most relevant opportunities in manufacturing industry derived from Data Economy?

The goal is to collect comments, ideas and personal opinions in order to:

- create a shared awareness and a profitable starting point about the main issues that are currently affecting the manufacturing domain,
- depict the current state of the art of Data Space in manufacturing in projects and research,
- identify the advantages of adopting Data Economy as an incentive to be showed to enterprises to drive them in the transformation.

In order to collect inputs and to benefit from a shared session where participants can be stimulated by the others' ideas, we have organized two interactive workshops: one with CF2 community, the other one with BOOST4.0 partners, both in the first half of December 2020.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.



Figure 30 – MURAL Interactive Workshop Board about DS Pathway

The picture above shows the interactive board used for the workshop: it consists of three different panels where the 6*5 evolutionary matrix is replicated. Focusing on columns (Data Space evolution levels) and taking inspiration from rows (areas impacted by Data Space evolution), participants were asked to reflect about:

- Obstacles and challenges, in the Blue Panel
- Success Stories, in the Green Panel
- Future Opportunities, in the Purple Panel

5.4 Validation and Experimentation of the Data Spaces Pathway

5.4.1 BOOST4.0 Experimentation

From the session held in December 2020 involving BOOST4.0 partners (pilot, replication factories and technology providers), it emerged that "**lack of trust in data sharing**" and "**complexity in data interpretability**" are the two most present barriers that block the data exploitation journey.

In large companies, moreover, a heterogeneous overview from the management, that doesn't share a common approach of how handle and exploit data potential, may add further complexity in reaching an agreement in the enterprise data policy. One suggestion to overcome this problem is the implementation of **PoCs** or draft solutions (quick but effective), to promote data power.

Another important aspect is the **lack of skills** regarding data modelling, rather than lack of infrastructures: data are available, as of now it is quite easy to collect, store and manage them, but it is not equally easy to implement and interpret AI models (which are the core of data exploitation).

Regarding opportunities for the future, from the point of view of the single company, an effective adoption of **data driven models** (deriving, of course, by a strong trust in them) could improve and assist decision making activities and increase the value of the final product sold to clients.

From the viewpoint of a wide network of enterprises, it is important that as many companies as possible are aware of data potential and start the journey toward data exploitation, since only a **shared awareness** may allow to create **Open Source solutions** and to create **Data Marketplaces**, increasing data value and opportunities.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

5.4.2 DMP Cluster Experimentation

In the workshop hold with DMP (Digital Manufacturing Platform) cluster community, it emerged some of the issues identified also discussing with BOOST4.0 partners. Since they are perceived as major obstacles by lot of participants, it is worth to highlight them again in order to stress the fact that they are barriers very well spread and rooted among enterprises. Comments pointed out: the issue of **lack of skills** regarding data modelling, the issue of lack of trust in data due to **low transparency** in the process of generating and sharing data, the need of **standard data model** to ease the interpretation of data and the need **PoCs** to spread awareness of the added value in using data.

Considering enterprises stuck to the "Data Siloed" level, the challenge is to provide a **good business approach to stakeholders** (for example, software vendors) to facilitate the use of data and also use more service based solutions.

Discussing with CF2 community, several ongoing (or just terminated) projects came out as example of research and investments in Data Space evolution, where new solutions to exploit data potential have been developed. For instance, in BOOST4.0 it has been implemented a solution for the warehouse optimization; in MIDIH for logistic optimization and predictive maintenance; in Market4.0 a Marketplace of Services and Equipment; in AI REGIO the goal is to create a network of DIH supporting manufacturing enterprises in AI implementation.

Regarding opportunities for the future, it clearly emerges the importance of creating a network of "Data - driven" enterprises, aware of data potential: reiterating what was express in BOOST4.0 workshop, it came out that data exploitation can be effective only if a proper Data Market is in place. It will allow to benefit from **Open Source solutions** and to use **data from different domains**.

According to more than a participant, after having reached the full exploitation of data, the step after "Data valorisation" in the Data Space pathway is exploring the world of **data from emotion**.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 873086.

6 Conclusions and Future Outlook

D2.2 has addressed and discussed a first challenge related to Digital Transition for Manufacturing, one of the two Twin Transitions of next Horizon Europe programme. The “Data Transition” has been analysed under different perspectives and an evolutionary Data Spaces pathway developed by crossing six Data Technology enablers versus a five-levels maturity model.

Preliminary validation of the model has been performed in the last months against some champion projects in the Digital Manufacturing Platform ecosystem. Such a validation will continue in the next few months also with a cross-sector perspective thanks to the OPEN DEI Digital Transformation CSA.

Next Digital Technology challenges, e.g. AI Cybersecurity HPC/cloud, will be addressed in the next period of the project and will be reported in the next edition of this deliverable.

The outcomes of this validation will be particularly useful for the implementation of Horizon Europe Destination 1 topics and in particular for the TWIN-TR-01-07: Artificial Intelligence for sustainable, agile manufacturing and the TWIN-TR-01-08: Data-driven Distributed Industrial Environments, where maturity levels could be used for a personalised adoption strategy of both Data and AI technologies.

On the other side, in the DEP “Data Spaces Support Centre” such CF2 methods and tools could be extensively adopted to the different multi-domain implementations of Data Spaces, the Manufacturing one in particular.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 873086.