D4.1 Reference Architecture for AI DIH

Author:	ENG
Work Package:	WP4 Beyond PLATFORMS: AI DIH Open Platforms and DIH Platform
Delivery date:	04.08.2021
Due date:	30.06.2021
Classification:	Public
Type:	Report

H2020 Innovation Action - This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N. 952003

Status of deliverable

Action/role	Name	Date (dd.mm.yyyy)
Submitted by	ENG	04.08.2021
Responsible (WP leader)	ENG	04.08.2021
Approved by (internal reviewer)	FZI	04.08.2021

Revision History

Date (dd.mm.yyyy)	Version	Author	Comments
10.05.2021	0.1	ENG	Deliverable 1st ToC
16.07.2021	0.2	ENG	1. RA backgrounds 2. RA specifications 3. Contributions from: • (CARTIF + SUITE5) • (TECNALIA + CEA) • (UNP) • (EXPERT-AI) • (CNR) • (EUT) • (NISSA)
21.07.2021	0.3	Contributions from Partners	Contributions from: • (CARTIF + SUITE5) • (TECNALIA + CEA) • (UNP) • (EXPERT-AI) • (CNR) • (EUT) • (NISSA)
23.07.2021	0.4	ENG	1. Internal review and add missing section 2. Contributions from: • (§2.2 UNP)
26.07.2021	0.5	ENG	Contribution from: • CARTIF (§3.1.3.3.2, §3.1.2) • CNR
27.07.2021	0.6	ENG	Contribution from: • SUITE5 (§3.1.2) • TECNALIA (§3.1.3.5.2) • NISSA (§3.1.1, §3.1.3.2.2)
29.07.2021	0.7	ENG	Contribution from: • CEA (§3.1.1, §3.1.3.5)

04.08.2021	1.0	ENG	Final version based on internal review comments (FZI)

Author(s) contact information

Name	Organisation	E-mail
Angelo Marguglio	ENG	angelo.marguglio@eng.it
Gabriele De Luca	ENG	gabriele.deluca@eng.it
Mattia Giuseppe Marzano	ENG	mattiagiuseppe.marzano@eng.it

Executive Summary (ENG)

The AI REGIO project aims to lowering the barriers preventing AI-driven DIHs from implementing fully effective digital transformation pathways for their Manufacturing SMEs.

In the DIH focus area, the need is to exploit existing open platforms coming from previous I4MS phases in relation to several domains (i.e., CPS/IOT, robotics, etc.) and targets (i.e., Data4AI, AI4Manufacturing, etc.), integrating new assets and solutions coming from SMEs, I4MS projects and the AI4EU initiative.

In order to transform DIHs in "honest brokers" of technology, in synergy with the other tasks of WP4, four aspects have been taken into account:

- an open source Data4AI platform, based on a data-driven approach to enable a real data economy in the domain of AI for manufacturing. It can be reached defining data preparation pipelines convenient even for non-technical expert, checking at the same time the completeness and validity of data.
- an AI4Manufacturing toolkit, a set of data analytics tools and techniques allowing the integration of different relevant outcomes in terms of algorithms, software frameworks, development tools and datasets.
- an AI-oriented trusted Data Sharing Space, where the data can be transferred in a secure way in order to achieve the desired outcome taking advantage from the knowledge and experience of the community.
- an Innovation/Collaboration platform for DIH to exchange the information with the perspective of increase the production quality and even the supply chain processes.

This deliverable addresses the definition of the AI REGIO Reference Architecture aligned to the state-of-the-art Industry 4.0 Reference Architectures and Reference Models from previous I4MS and AI-related initiatives. Starting from the BDVA Grand Challenges and requirements collected in the D2.3 and D2.5, the design phase has considered several scenarios coming from real needs described in the trials.

The proposed solution aims to support the manufacturing scenarios where the adoption of the Artificial Intelligence implements the "AI-transformation" into an ecosystem (of factories) domain, integrating the Data Space building blocks as described in the recent position paper [3].

Furthermore, a Reference Implementation, based on open-source components, is defined in order to lower the barriers for the adopters, acting as a baseline for the experiment deployments.

The document has been elaborated by the main partners involved in the WP4, in strict synergy with the WP4 task leaders.

Contents

Figures

Tables

1 Introduction

1.1 Scope of the Deliverable

In the context of Industry 4.0, the Artificial Intelligence can help companies to be more productive and efficient. According to the Deloitte Survey on AI Adoption in Manufacturing, 87% of manufacturing companies have adopted AI or planned to do so within 2022 and 83% hold that AI will make a tangible impact on manufacturing and management by 2025 [1].

Artificial Intelligence is one of the disruptive technologies that are transforming business especially during a crisis. The pandemic, indeed, has demonstrated how AI can do in so many different ways.

Furthermore, the digital transformation in the manufacturing market is on the road and it will increase again in the next years thanks to enabling technologies like IoT, Robotics, 3D Printing, etc. On the other hand, the manufacturing sector has one of the lower indicators (43%) for data-driven decision making [2].

The AI REGIO project aims to fill this gap, defining a Reference Architecture for describing a wide range of industrial scenarios, covering the entire data flow with the integration the AI-driven data processing and empowering the data economy. Considering the existing solutions as outcomes of the previous I4MS phases, Artificial Intelligence will impact existing processes bringing innovation exploiting available data in order to extract information and support both blue and white collars. Supporting the Data Sovereignty principles, instead, is the first mandatory step for enabling the data exchange at the ecosystem level, exploiting the related advantages for data providers and consumers applying the governance rules and technical solutions for secure and sovereign data exchange.

The analysis of the Industry 4.0 state-of-the-art in terms of reference architectures, standards and protocols will represent the main input for the definition of the AI REGIO Reference Architecture. Starting from it, a Reference Implementation will be defined, composed of open-source components, in order to enable the AI adoption in synergy with the DIH activities.

1.2 Contributions to other WPs and Deliverables

AI REGIO project aims to fill the three major gaps currently preventing AI-driven DIHs from implementing the fully digital transformation for the manufacturing SMEs. The key strategy is to align the VANGUARD innovation strategy with the outcomes of the I4MS phases. Leveraging existing assets, complete open platforms and marketplaces, AI REGIO will address the need for Artificial Intelligence, addressing the AI Industry 5.0 challenges in relation to the human-machine interaction and human acceptance. The digital transformation of SMEs can be addressed following dedicated programs supported by DIHs in terms of competencies acting as facilitators.

The project will define an open Reference Architecture for AI scenarios totally implementable with open-source components. The initial requirements will be collected starting from the three challenges defined in the BDVA SRIA [4] paper:

- Smart Factory: data is generated inside production lines and analytics is needed for safety, optimization and diagnosis of the plant as well as of the workers.

- Smart Product: data is generated by ecosystems of suppliers, providers, distributors, retailers and analytics is needed for value chain integration and workers collaboration.
- Smart Supply Chain: data is generated by the product-service itself along its lifecycle in Circular Economy perspective and analytics is needed for product operations monitoring and control.

In order to emphasize the Data Economy aspect, a Cross-Sector challenge is also considered, covering the need for communication between companies (and factories).

Furthermore, requirements collected in the context of WP2 and detailed into D2.3 and D2.5, will represent the input for the refinement phase fitting the needs of all AI REGIO clusters composed of 17 experiments for implementing different use cases.

The Reference Architecture and the proposed Reference Implementation will be validated in the context of WP6 during the implementation of the use cases.

1.3 Structure of the Document

The rest of the deliverable is structured as follows:

- Section 2: backgrounds in terms of reference architectures, briefly described, to design AI REGIO RA in the following section. Standards and protocols useful to develop the platform properly.
- Section 3: AI REGIO Reference architecture designed and structured following the design principles and having as milestones the 5 sub-systems (Industry 5.0, Data4AI, AI4Manufacturing Toolkit, DIHIWARE, AI Data Space for Manufacturing). Definition of Reference Implementation empowering the already know tools for Industry 4.0 with AI tools.
- Section 4: description of the achieved results and next steps can be reached from this checkpoint.

2 AI REGIO Reference Architecture Backgrounds

2.1 Backgrounds and relevant initiatives

In this Section the main solutions and initiatives will be described, paying a particular attention on the relevant Reference Architectures in relation to data-driven manufacturing. These will represent the baseline for designing the AI REGIO Reference Architecture.

The AI REGIO project aims to define a Reference Architecture for describing a large variety of industrial scenarios, where Artificial Intelligence will have effects on the existing processes. Data Sovereignty principles will be one of the main parts of the architecture to exchange data in the ecosystem. For this purpose, will be used the IDSA Connectors, including even all the connected services related (e.g., IDS Metadata Broker, IDS Identity Provider, Clearing House).

Currently, the Industrial Internet Consortium (IIC) and the Working Group for Industry 4.0, have been working to provide recommendations and guidelines following the structure of the Industrial Internet Reference Architecture (IIRA) and the Reference Architecture Model for Industry 4.0 (RAMI 4.0), respectively.

Thus, from these starting points, the objective of this document is to give a data-driven approach for its reference architecture capable of driving the industrial sector transformation into a networked, data driven environment, empowered at the same time with the concepts of AI that will supports the entire scenario, adopting several patterns for gateways, edge architectures and data buses at different layers.

Open initiatives such as FIWARE and additional open-source components will be able to support the AI REGIO platform, defining high value applications for manufacturers to optimize production systems and value chains, and radically improve sustainability, productivity, innovation and customer service using the advances made in big data and data analytics through both, the data-in-motion and the data-at-rest.

Architectural IIoT systems where connected machines, or industrial systems, exchanging data, analysis, and actions to improve production and performance, and reduce defects or failures, are the main features of the framework. A huge amount of data is collected and analyzed, putting them at the disposal of the final user who is in charge of work with the tools allowing the factories to upgrade their performance.

2.1.1 Reference Architectural Model Industrie (RAMI 4.0)

RAMI 4.0 (Reference Architectural Model Industrie 4.0) [6] purpose's is providing a reference model solution to the applications that use Internet of Things (IoT), big data analytics, and other technologies advancements in manufacturing processes. It is based on a three-dimensional model and organizes the manufacturing hierarchy levels across the six layers of the IT representation of Industry 4.0.

A three-dimensional matrix is used to describe standards and use-cases. It addresses integration within and between factories, end-to-end engineering and human value-stream orchestration.

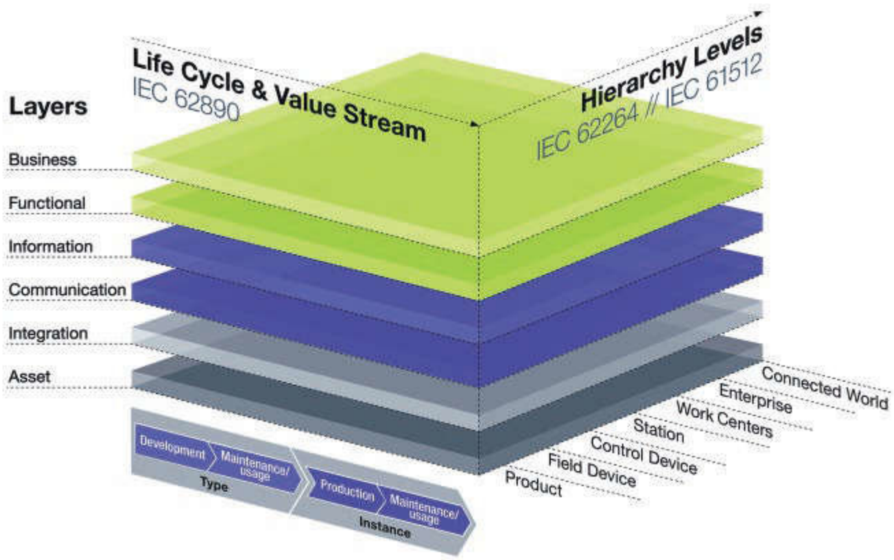


Figure 1 - The three dimensions of the RAMI 4.0 (Source: Platform I4.0 And ZVEI)

In RAMI4.0, as mentioned before, each component consists of six layers showed in the figure below, where:

- The **asset layer** describes the physical components of a system.
- The **integration layer** is useful to compute information contents and is a kind of link between the real and the IT world. It contains field buses, HMIs, all the needs to use an asset and generate events based on the retrieved information.
- The **communication layer** standardizes communication between integration and information layer: transmitting data and files, flattening data formats, protocols and interfaces.
- The **information layer** has data in a structured form and gives the access to the structured data from. It processes, integrates and persists the data and events, as well as describes the data related to the technical functionality of an asset.
- The **functional layer** defines logically and technically an asset, describing its functions and a platform for horizontal integration, the business model mapping, business processes adjustable on inputs from the functional layer.
- The **business layer** holds the orchestration of services provided by the functional layer. It connects the services to the business models and the business process models.

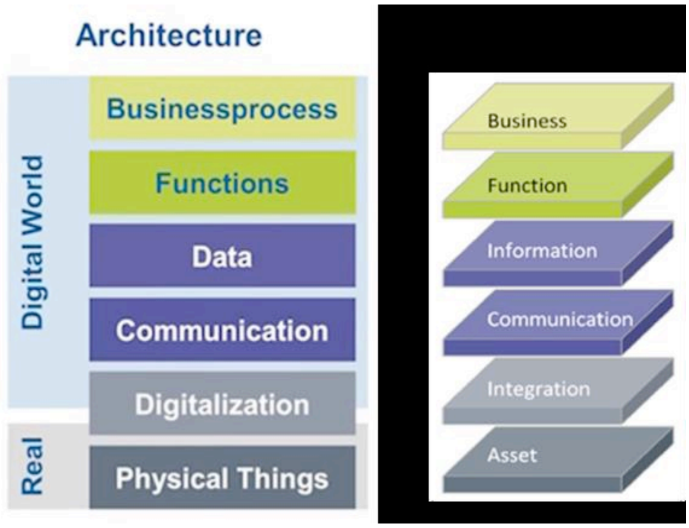


Figure 2 - The IT Layers of RAMI 4.0

Asset Administration Shell

The Asset Administration Shell (AAS) is the common representation for the integration of the physical assets. Its goal is bridging real world and the IoT world integrating assets into the world of information.

An asset (i.e., machinery, parts, supply material, documents, contracts, etc.) defines the data models for exchanging information between partners in the value chain and a package file format (the Asset Administration Shell Package, AASX), to exchange the full or partial structure of the administration shell.

The AAS ensures interoperability, availability, Integration, and can be provided in many implementation ways. Starting from the classification given, it relies to the RAMI4.0 Model covering different layers.

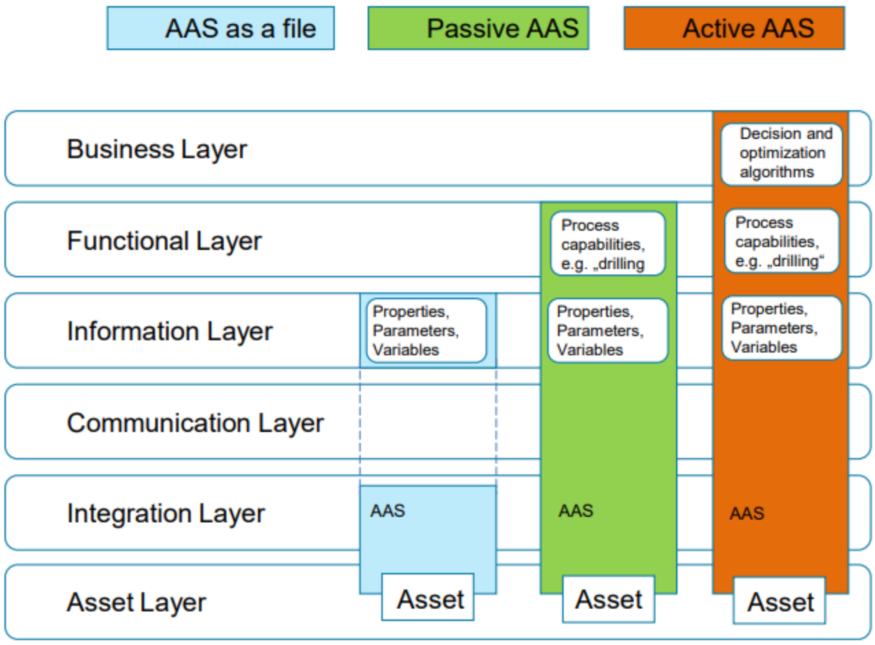


Figure 3 - Assignment of the active AAS in the RAMI4.0-Model

2.1.2 Industrial Internet Reference Architecture (IIRA)

The Industrial Internet Reference Architecture (IIRA) [7] tries to identify the most important and common architecture concerns. For this scope it creates an architectural template and methodology that can be used to examine and resolve design issues. Furthermore, it allows designers to get insights by examining architecture patterns and help them to avoid missing relevant architecture and design gaps.

The IIRA puts the accent on interoperability and practical deployment of IoT assets. It is not used to specify a low-level details development of a system, but is mainly used for specifying principles of IoT systems, as well as for communicating concepts and boosting stakeholders' collaboration.

The IIRA approach concerns a set of frames called viewpoints that allow designers to discover and find a solution to the main issues. These viewpoints are sorted in order to represents the interactions between them, because the choice of a higher-level viewpoint must require the viewpoints below.

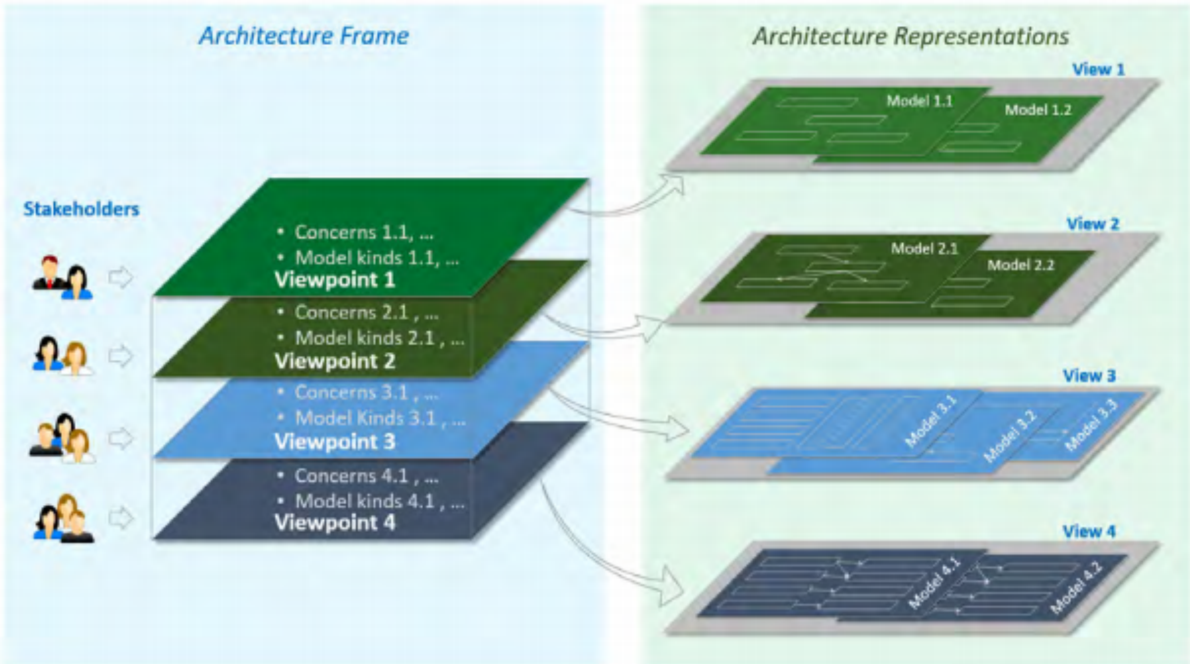


Figure 4 - IIRA Architecture Framework (IIC, 2019)

In the figure above there are the 4 viewpoints making IIRA a layered architecture, in the following section they are explained more in detail:

- **Business viewpoint** takes in account the stakeholders involved in the development, deployment and operation of an IoT system.
- **Usage viewpoint** considers the usage of the IoT system like a sequence of activities that may be performed by actors.
- **Functional viewpoint** specifies the functionalities of the system, also presenting interactions with external logical modules.

- **Implementation viewpoint** comprises the development technologies used to make the components, along with information about their lifecycle and the realization of the communication between them.

The most relevant is the functional viewpoint, because it specifies a number of different functionalities, which are called functional domains. In this way, ideally an IoT system can be fragmented into "functional domains", which are the building blocks applicable to domains and applications.

2.1.3 International Data Spaces Reference Architecture Model (IDS-RAM)

The international data spaces (IDS) is a virtual data space to facilitate secure and standardized data exchange and data linkage in a trusted business ecosystem. It allows the designer to have a basis to make smart-service scenarios and easily innovative cross-company business processes, at the same time ensuring data sovereignty for data owners.

The IDS-RAM [5] is at a higher abstraction level than other architecture models, actually gives an overview about individual components of the International Data Spaces (Connector, Broker, App Store, etc.). In detail, it is composed by five-layer structure, where each layer concerns stakeholders' viewpoints at different levels of granularity, as showed in the figure below.

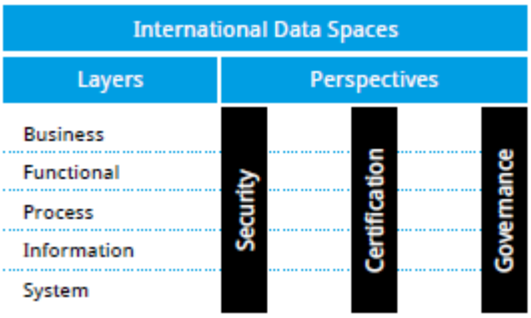


Figure 5 - General structure of Reference Architecture Model

- The **Business Layer** classifies the different roles can have the participants of the International Data Spaces with the main activities and interactions with each of these roles.
- The **Functional Layer** defines the functional requirements of the International Data Spaces, plus the concrete features to be derived from these.
- The **Process Layer** is related to the interactions between the different components of the International Data Spaces.
- The **Information Layer** is a conceptual model which uses linked-data principles to describe the static and the dynamic aspects of the International Data Space's constituents.
- The **System Layer** takes into account the decomposition of the logical software components, considering aspects such as integration, configuration, deployment, and extensibility of these components.

Furthermore, the Reference Architecture Model embraces three perspectives that need to be implemented across all five layers:

- **Security** is a strategic requirement of the International Data Spaces to ensure secure data supply chains. The main goal is maintaining the trust among Participants that want to exchange and share data and use Data Apps. To reach that, the architecture has inside a Trusted Connector which is in charge to materialize the Security Architecture in everyday interactions and operations in the International Data Spaces.
- **Certification**: Data security and data sovereignty are the pillars of International Data Spaces, where Data sovereignty can be defined as a natural person's or legal entity's capability to be in full control of its data. While the certification is focused on security and trust, the certification of components also refers to compliance with technical requirements ensuring interoperability.
- **Governance**: The Governance Perspective describes the roles, functions, and processes of the International Data Spaces and defines the requirements needed in the business ecosystem to achieve secure and reliable corporate interoperability.

The following figure shows a typical architecture stack of the digital industrial enterprise, that connects the lower-level architectures for communication and basic data services with more abstract architectures for smart data services. It therefore supports the establishment of secure data supply chains from data source to data use, while at the same time making sure data sovereignty is guaranteed for data owners.

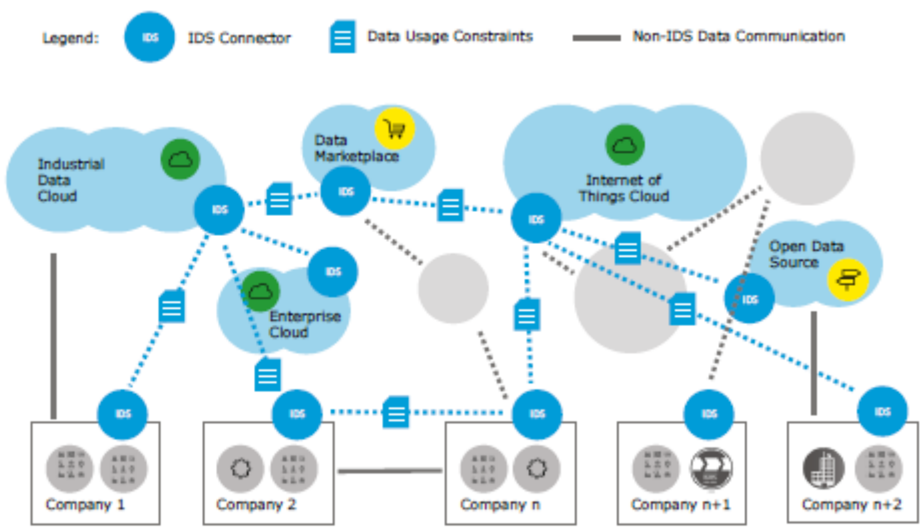


Figure 6 - International Data Spaces connecting different cloud platforms.

2.1.4 Data, AI and Robotics (BDVA/DAIRO)

The BDV Reference Model [4] has been developed by the BDVA, taking into account input from technical experts and stakeholders along the whole Big Data Value chain, as well as interactions with other related PPPs. The BDV Reference Model may serve as common reference framework to locate Big Data technologies on the overall IT stack. It addresses the main concerns and aspects to be considered for Big Data Value systems. The BDV Reference Model distinguishes between two different elements. On the one hand, it describes the elements that are at the core of the BDVA; on the other, it outlines the features that are developed in strong collaboration with related European activities.

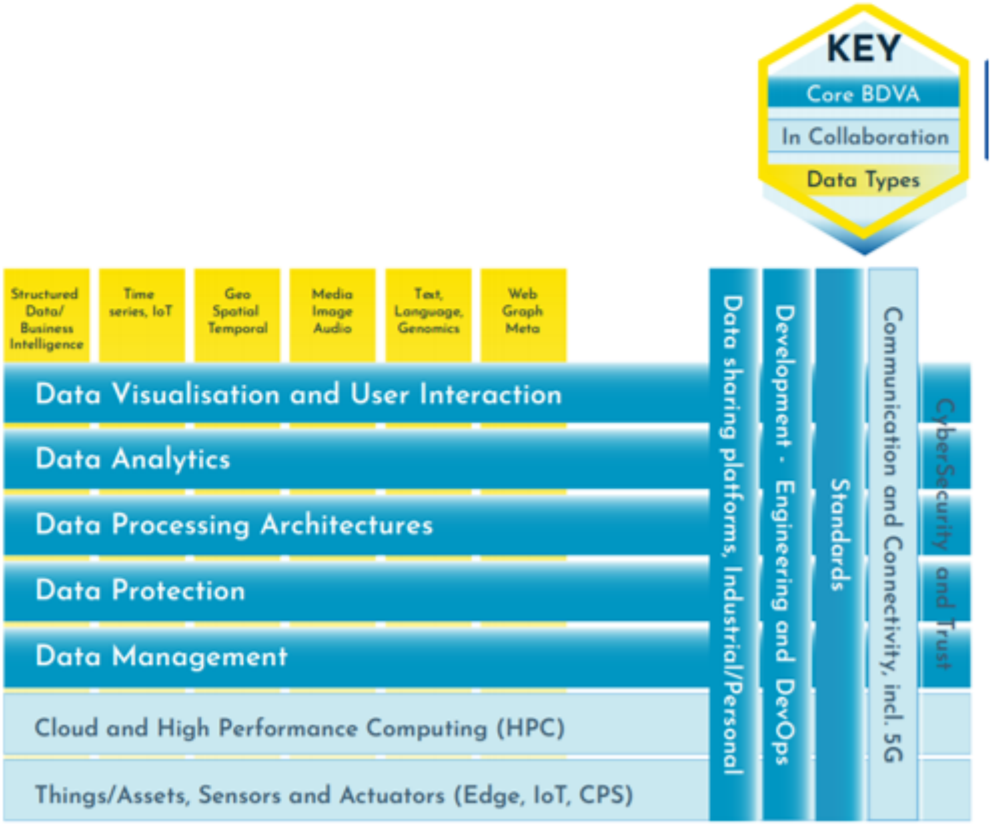


Figure 7 - Big Data Value Reference Model

The BDV Reference Model is structured into horizontal and vertical concerns.

Horizontal concerns cover specific aspects along the data processing chain, starting with data collection and ingestion, and extending to data visualization. Briefly:

- Data Visualization and User Interaction: Advanced visualization approaches for improved user experience
- Data Analytics: Data analytics to improve data understanding, deep learning and the meaningfulness of data.
- Data Processing Architectures: Optimized and scalable architectures for analytics of both data-at-rest and data-in-motion, with low latency delivering real-time analytics.
- Data Protection: Privacy and anonymization mechanisms to facilitate data protection.
- Data Management: Principles and techniques for data management.
- The Cloud and High-Performance Computing (HPC): Effective Big Data processing and data management might imply the effective usage of Cloud and High-Performance Computing infrastructures
- IoT, CPS, Edge and Fog Computing: A main source of Big Data is sensor data from an IoT context and actuator interaction in Cyber Physical Systems.

Vertical concerns address cross-cutting issues, which may affect all the horizontal concerns. In addition, vertical concerns may also involve non-technical aspects. Briefly:

- Big Data Types and Semantics: The following 6 Big Data types have been identified, based on the fact that they often lead to the use of different techniques and mechanisms in the horizontal concerns, which should be considered, for instance, for data analytics and data storage: (1) Structured data; (2) Time series data; (3) Geospatial data; (4) Media, Image, Video and Audio data; (5) Text data, including Natural Language Processing data and Genomics representations; and (6) Graph data, Network/Web data and Metadata. In addition, it is important to support both the syntactical and semantic aspects of data for all Big Data types

- Standards: Standardization of Big Data technology areas to facilitate data integration, sharing and interoperability.
- Communication and Connectivity: Effective communication and connectivity mechanisms are necessary in providing support for Big Data
- Cybersecurity: Big Data often need support to maintain security and trust beyond privacy and anonymization.
- Engineering and DevOps for building Big Data Value systems
- Marketplaces, Industrial Data Platforms and Personal Data Platforms (IDPs/PDPs), Ecosystems for Data Sharing and Innovation Support: Data platforms for data sharing include, in particular, IDPs and PDPs, but also other data sharing platforms like Research Data Platforms (RDPs) and Urban/City Data Platforms (UDPs).

2.1.5 Artificial Intelligence for Manufacturing SoTA

AI is one of the most prominent and most disruptive digital technologies for Industry 4.0 systems and applications. Some of the most prominent use cases include:

- Predictive maintenance and intelligent asset management, where AI algorithms such as Deep Learning (DL) are used to predict and anticipate asset parameters like Remaining Useful Life (RUL) and End of Life (EoL), in order to schedule maintenance, service and repair activities at the best point in time.
- Predictive quality management (Quality 4.0) and zero-defect manufacturing (ZDM), which process large amounts of digital data from a production line by means of AI algorithms in order to understand and anticipate situations (e.g., process configurations, production parameters) that will likely to lead to defects. Leveraging knowledge about these situations AI systems are developed to avoid or remedy these situations.
- Generative product design, which makes use of AI algorithms to produce many and different product variants towards maximizing/optimizing parameters like efficiency, sustainability and cost.
- Human Robot Collaboration, where AI systems are used to develop, deploy and operate co-bots. The latter delegated tedious, laborious and repetitive processes to robots, while ensuring their proper collaboration with humans that undertake other parts of the production process.

2.1.6 FIWARE Smart Industry Reference Architecture (F4I)

FIWARE[1] is framework of open-source platform components which can be assembled together and with other components to facilitate the development of interoperable and portable solutions in many domains.

FIWARE manages the data by supporting access to a Context / Digital Twin representation, using FIWARE NGSI, that is the RESTful API, to publish and access data. This is implemented by the

[1] FIWARE developers catalogue. (n.d.). Retrieved from https://www.fiware.org/developers/catalogue/

Context Broker, one of the main components of FIWARE architecture.

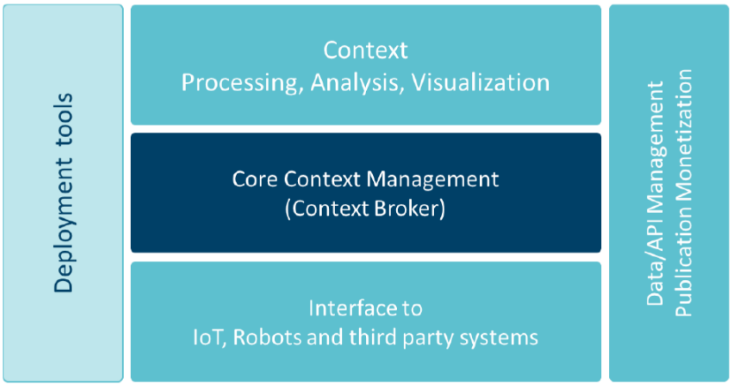


Figure 8 - FIWARE overall architecture

FIWARE GEs are structured like in the figure above, where FIWARE Context Broker Generic Enabler is the only mandatory component of all solutions.

Starting the development around the FIWARE Context Broker, the designer can enrich the solution with many types of FIWARE components available:

- **Interfacing with the Internet of Things (IoT), Robots and 3rd Party systems**, for capturing updates on context information and translating required actuations.
- **Context Data/API management, publication, and monetization**, bringing support to usage control and the opportunity to publish and monetize part of managed context data.
- **Processing, analysis, and visualization of context information** implementing the expected smart behavior of applications and/or assisting end users in making smart decisions.

The FIWARE catalogue contains a wide library of enablers with reference implementations that help developers to put in place many functionalities, making easy the programming, facilitating at the same time IaaS and SaaS application domains.

Smart Industry

FIWARE for Industry (F4I) is a multi-project initiative with the intent of develop an ecosystem of FIWARE-enabled software components, to help the needs of Manufacturing Industry business scenarios. F4I in 2015 developed open-source reference implementations of Smart-Digital-Virtual Factory scenarios with the integration of 14 FIWARE Generic Enablers and 15 Manufacturing Industry Specific Enablers.

Starting from these 29 components, more than 15 new projects are developing their Smart Manufacturing solutions in several R&I domains. In the Factories of the Future H2020 cPPP, the recent projects are contributing in kind to the picture here below which includes 17 Enablers: 8 enhancements of FITMAN SEs, 3 enhancements of FIWARE GEs, 2 new specific enablers and 4 new releases of FIWARE GEs. Moreover, the EIT DIGITAL High Impact Initiative called OEDIPUS (Operate European Digital Industry with Products and Services) is developing platforms and components FIWARE based for the Smart Manufacturing Industry, in close collaboration with SIEMENS (OEDIPUS coordinator) and its platforms (e.g. MindSphere). In the near future, the EU-Brasil FASTEN project and some National / Regional projects will give their contributions as well to the F4I ecosystem.

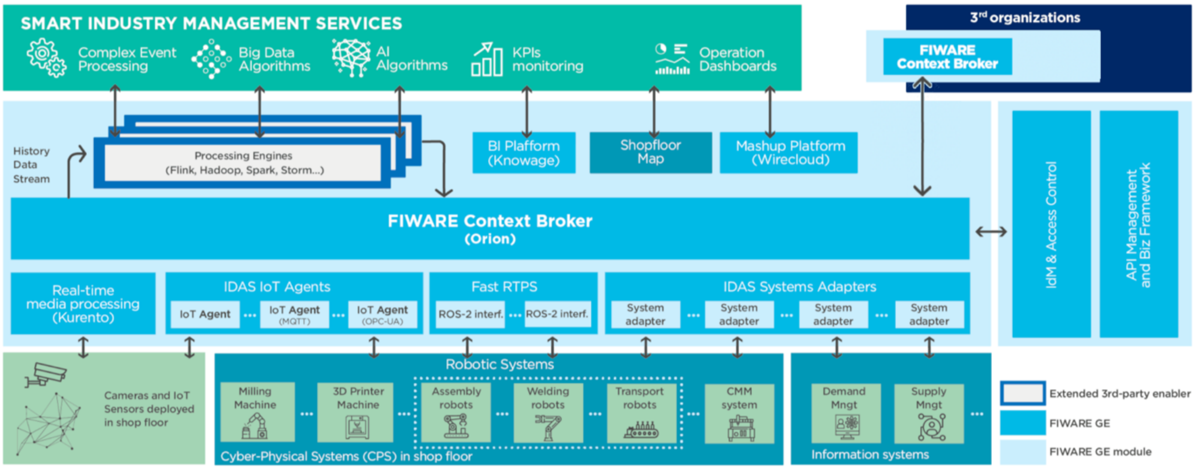


Figure 9 - Reference Architecture for Smart Industry Management System

The FIWARE Foundation Smart Manufacturing developed a Reference Architecture, as reported below, which is the common top-down basis for building successful applications on top of both FIWARE GEs and F4I SEs. Following a data-driven approach, the F4I RA is based on the Data in Motion (Industrial IOT) and the Data at Rest (Industrial Analytics) processing modality.

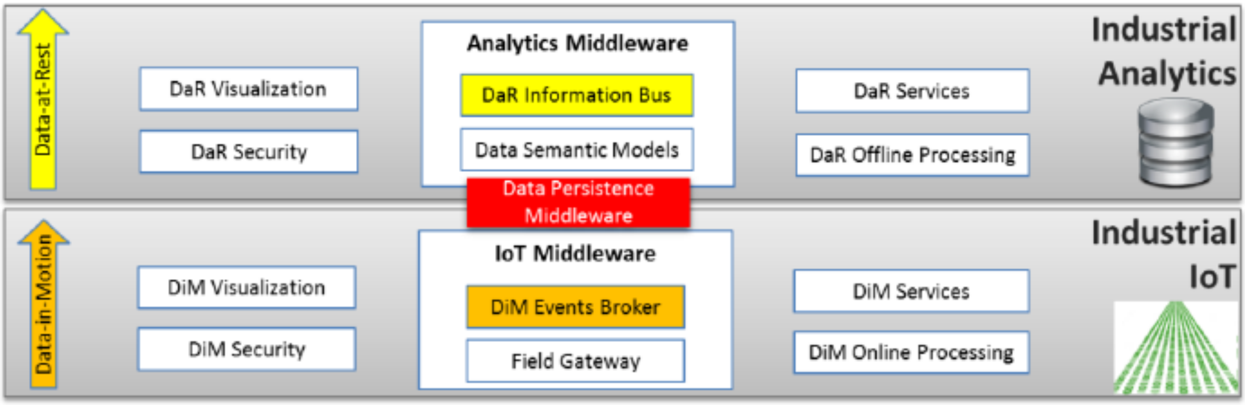


Figure 10 - Draft of a data-driven Reference Architecture

2.1.7 Manufacturing Industry Digital Innovation Hubs (MIDIH)

The MIDIH [8] Reference Architecture has the goal to identify and define at high-level the functionalities basing itself on a meta-model where the four main RAs could be aligned and interoperated for the development of Smart Factory, Smart Product and Smart Supply Chain reference implementations: the FIWARE RA, the IDS RA, the IIC RA and the Platform Industrie 4.0 RAMI 4.0 have been therefore considered, as main inputs for MIDIH. It mainly takes inspiration from FIWARE for Industry solution, at the same time a data-driven approach is followed, enhanced by Industrial IOT and Industrial Analytics functions. For this scope, the core of the MIDIH RA will cover a lot of analytics framework interconnected having as central features **data in Motion,** which is produced by many real assets being them in the Smart Factory or in the Smart Product, and **data at Rest,** that need to be processed in order to feed AI-based advanced applications in different fields of Factory.

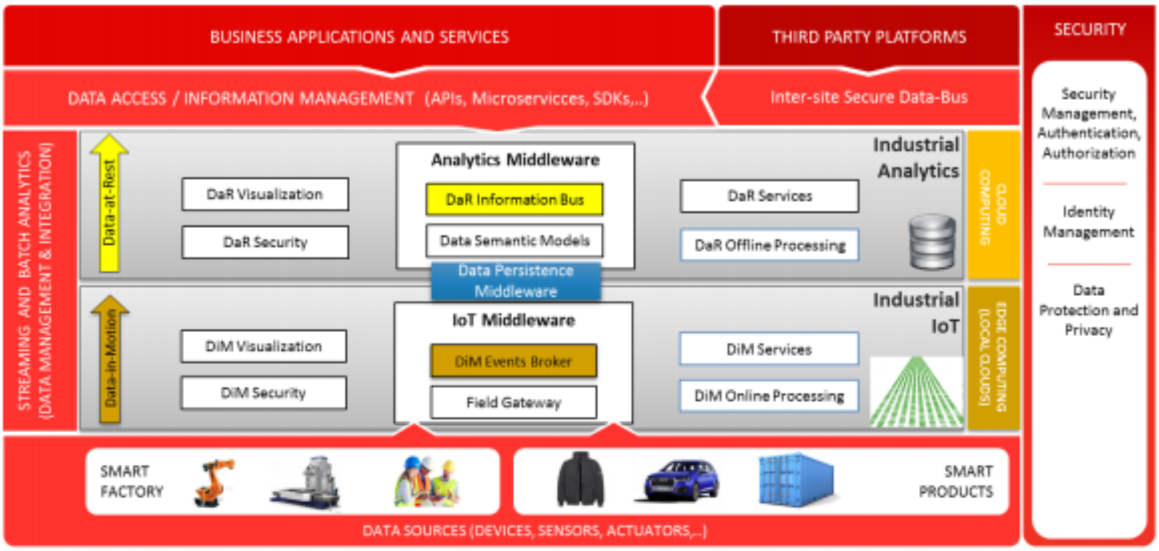


Figure 11 - MIDIH Reference Architecture

To complete the scenario, at the bottom of RA there is the real-world assets, while at the top is identified the Application Ecosystem.

The Layered Databus pattern, in the instance of the three tier architecture was the starting point to build and design the RA:

- **Field or Edge Tier** is mapped with the Unit Databus.
- **Platform and Enterprise Tiers** are mapped with the Site Databus.
- **Business Ecosystem Tier** is a new layer entrusted to the three-tier architecture pattern to enable inter-factory communications that will allow the creation of a reliable MIDIH business ecosystem, and it is mapped with the Inter-Site Databus.

In parallel security along the different layers will be taken into account. The requirements coming from the different experiments specified in the three different scenarios: Smart Factory, Smart product and Smart Supply chain have contributed to define some of the functional components.

2.2 Standards and Protocols for Industry 4.0

When talking about IoT, it is always done a link with communication, and interaction between hardware devices (sensors, gateways, servers). What makes this communication between devices possible are the IoT protocols.

IoT protocols are an important part of IoT technology as they enable hardware to exchange data in a structured and meaningful way.

In this section, it is looked into some of the IoT protocols used in the industry, namely, Industrial Connectivity Protocol, OPC UA, MQTT, CoAP, DDS and oneM2M.

2.2.1 Industrial Connectivity Protocol

Data acquisition and data sharing are two of the most important functions of any industrial AI system: AI systems are data driven and need (historical) industrial data for training models and algorithms, as well as fresh data to perform their tasks like classification or extraction of associations. Therefore, connectivity functionalities are integral elements of AI deployments. These functionalities enable different components of Industrial Internet of Things (IIoT) and Industry 4.0 systems to connect to data sources for the purposes of acquiring industrial data as needed for their operations.

State of the art IIoT systems leverage a very rich set of connectivity technologies and standards in order to connect to different industrial data sources like sensors, automation devices and Cyber Physical Production Systems (CPPS). Most of these technologies are optimized for specific applications or use cases in vertical industries. Nevertheless, most of them are not designed in ways that facilitate data interoperability across different data producers and consumers, and do not cater for sharing data across consumers. This is a setback to supporting the interoperability across legacy components (e.g., brownfield subsystems) and new Industry 4.0 components (i.e., greenfield subsystems), but also within different stakeholders in the manufacturing chain. In order to alleviate this situation, industrial organizations should with standards-based connectivity technologies that span different connectivity layers i.e. from the physical and network layers to the network and transport layers. Most importantly, they must leverage standards-based industrial protocols that boost data sharing and data interoperability across IIoT components and applications. Following paragraphs outlines some of the most prominent industrial standards for connectivity, data sharing and data interoperability in the context of Industry4.0 applications. The listed standards and frameworks are also listed in the scope of the Industrial Internet Connectivity Framework[2] as being appropriate to support reliable and interoperable connectivity in-line with the requirements of the Industrial Internet Reference Architecture (IIRA). AI REGIO takes into account these standards and caters for supporting their operation as part of the project's reference architecture and related "stack" of AI applications. It should also provide support, training and insights to SMEs and other stakeholders that would like to adopt and use them in their applications.

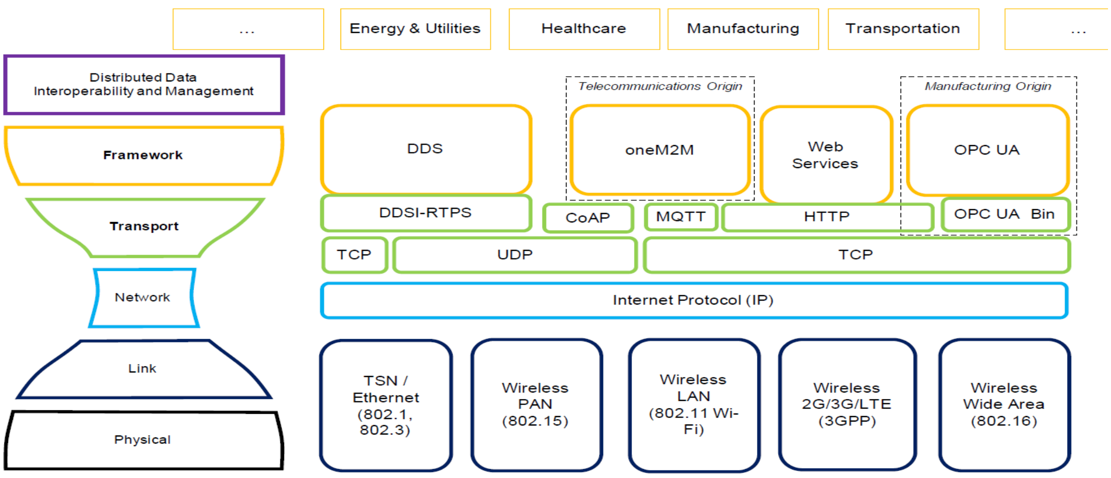


Figure 12 - Taxonomy of IIoT Connectivity Standards according to the Industrial Internet Connectivity Framework (IICF)

[2] Industrial Internet Connectivity Framework: https://www.iiconsortium.org/IICF.htm

2.2.2 OPC UA

OPC (Open Platform Communication) is the interoperability standard for the secure and reliable exchange of data in the industrial automation space and in other industries. It is platform independent and ensures the seamless flow of information among devices from multiple vendors.

The OPC standard is a series of specifications developed by industry vendors, end-users and software developers. These specifications define the interface between Clients and Servers, as well as Servers and Servers, including access to real-time data, monitoring of alarms and events, access to historical data and other applications. (What is OPC?, n.d.)

OPC UA is a platform-independent, service-oriented architecture specification that integrates all functionality from the existing OPC Classic specifications, providing a migration path to a more secure and scalable solution. (Unified Architecture, s.d.)

One of the biggest challenges faced by Industry 4.0 is Interoperability. OPC UA can address the Interoperability challenges that block the take-off of Industry 4.0.

OPC UA is based on an extensible framework with a multilayered architecture to accomplish these design goals:

1. **Functional Equivalence:** Build on top of OPC Classic specification
2. **Platform Independence:** Small embedded microcontrollers to cloud-based deployments
3. **Secure Communication:** Security built in from the ground up and firewall friendly
4. **Extensible design:** Add new features without affecting existing applications
5. **Comprehensive Information Modeling:** Industry vertical interoperability requirements (Unified Architecture, s.d.)

2.2.3 MQTT

MQTT is an OASIS standard for IoT connectivity. It is a publish/subscribe, extremely simple and lightweight messaging protocol, designed for constrained devices and low-bandwidth, high-latency or unreliable networks. The design principles are to minimize network bandwidth and device resource requirements whilst also attempting to ensure reliability and some degree of assurance of delivery. These principles also turn out to make the protocol ideal of the "Internet of Things" world of connected devices, and for mobile applications where bandwidth and battery power are at a premium. (MQTT FAQs, s.d.)

Several features of the protocol make it particularly suitable for remote sensing and control. Unlike the usual poll/response model that generally saturates data connections with unchanging data, MQTT maximizes available bandwidth. Update rates in the 100-millisecond range are possible even with external cloud-based brokers.

Being a lightweight and efficient protocol facilitates a higher throughput rate, which in turn significantly increases the amount of data that can be monitored or controlled. Therefore, organizations can publish stranded intelligence from field devices, such as flowmeters, to the MQTT server, and maintenance folks can subscribe to it (whereas operational clients would subscribe to operational data). Previously, this metadata had to be manually retrieved from the location, because it was often so voluminous that bandwidth limitations made transporting it out of the field hard to justify. (Hechtman, s.d.)

2.2.4 Constrained Application Protocol (CoAP)

CoAP is a connectivity transport standard based on HTTP, which (like HTTP) has been introduced and maintained by the IETF (Internet Engineering Task Force) as an open standard specification. It resides at the same level of the connectivity stack (i.e., transport layer) as MQTT and HTTP. Among its main merits is that it is more lightweight and better performing than HTTP. The IETF maintains the CoAP open standard specification. It is used to support operations domain and operates like HTTP: Clients send requests to a server, including information about a data object, an operation (e.g., PUT, GET, CREATE), and a payload. The server responds indicate failure or success of the operation, while providing a proper response payload. Subscriptions for notifications on the status of a data object are also supported. CoAP can be used for device-to-device communications, while supporting interoperable interactions with IIoT component that support HTTP and RESTful web services.

2.2.5 DDS

The DDS is an open connectivity framework for Industry 4.0 applications, which is developed and maintained by the Object Management Group (OMG)[3]. It is destined to connect IIoT components (e.g., devices, sensors, gateways or entire applications) to other components with a view to supporting real-time interactions and the creation of large systems as part of a Systems-of-Systems (SoS) approach. DDS components interact through a shared data space, which bears commonalities with the concept of Industrial Data Spaces (IDS). DDS components do not interact with each other.

DDS is commonly called "databus" as it manages data-in-motion (i.e. streaming data) much in the same way conventional databases manage data-at-rest. In this direction, the databus implements "data-centric" abstraction that is reflected on a relational model. DDS applications interact with the DDS infrastructure rather than directly based on P2P (point-to-point) interactions. Therefore, it supports loose coupling and independence of IIoT components. Most importantly, the databus provides functions for the dynamic discovery of diverse IIoT components prior to their interactions, which boost scalability, configurability and operation in dynamic and volatile environments. Likewise, a databus provides the means for implementing (semi)autonomous systems that can dynamically discover and use their data sources without human intervention.

A databus provides the means for filtering the incoming data, which is one of the main data management functions of DDS. This enables the implementation of scalable IIoT systems that integrate diverse IIoT components. Specifically, in order to add a new component, it is sufficient to integrate it within the databus. Once this takes place, the component can automatically discover and interact with other elements that are also connected in the bus, without any need for implementing costly and tedious P2P connections.

[3] Information about the DDS standard is provided at the DDS portal that is maintained by the DDS foundation: https://www.omg.org/omg-dds-portal/

2.2.6 oneM2M

oneM2M is managed by a partnership of regional international standards organizations, including ETSI. It provides a set of common service functions that operate between industrial applications and the transport layer of industrial connectivity protocols. These functions are accessible via RESTful APIs (Application Programing Interfaces). Using oneM2M industrial applications on different connected machines and devices become able to communicate with each other in secure and high-performance ways. oneM2M provides scalability and versatility as its functional/middleware elements can be flexibly deployed at the cloud, at the edge of the network or even on hosts and devices.

One of the main value propositions of oneM2M is that is can interface and support a variety of transport layer protocols, including HTTP, CoAP, MQTT and WebSockets, while there have been also efforts for supporting internetworking and interoperability between: (i) oneM2M and DDS and (ii) oneM2M and OPC-UA.

2.3 Standardization and Certification Services

In the context of the DIH services of AI REGIO, the project could offer standardization and certification services as follows:

- **List and searchable catalogue of AI standards for SMEs**: In this direction a range of AI standards with direct or indirect relevance to Industry 4.0 will be collected and made available through the AI REGIO DIH.
- **AI Maturity Assessment**: In the context of WP4, AI REGIO is introducing a new maturity model for manufacturing enterprises (including SMEs). The Maturity Model considers different activities e.g., Awareness, Active, Operational, Systemic, Transformational) and levels (e.g. AI Novice, AI Ready, AI Proficient, AI Advanced) of maturity of an enterprise in terms of AI adoption and use[4]. Likewise, different aspects (e.g., Strategic, Technology, Organizational, Regulatory etc.) are considered. AI REGIO is considering the development of an AI maturity assessment and certification scorecard as part of its DIH.

[4] T. Pringle and E. Zoller, "How to Achieve AI Maturity and Why It Matters An AI maturity assessment model and road map for CSPs," Ovum, 2018.

3 AI REGIO Reference Architecture

3.1 Architecture Requirements and Specifications

In this section requirements and specifications coming from the AI REGIO experiments, Artificial Intelligence challenges and AI REGIO Techno Assets will be analyzed. In particular, inputs from experiments will be listed, in synergy with the results of WP2 activities, describing the meaning and the context. The Artificial Intelligence challenges will complement specifications collected from the clusters, addressing generic AI-driven needs to be addressed by the AI REGIO Reference Architecture. Finally, the 5 AI REGIO sub-systems will be described including the related background, the main requirements and a conceptual architectural description for each one.

3.1.1 Inputs from Experiments

Cluster 1: Product Engineering and Lifecycle Management

Table 1 - Requirements for AI REGIO Platform based on Cluster 1 needs

Identifier	High Level Requirement	Description
CL.01.EXP.15.IN.01	*Security & Privacy*	*Once the API for integration into the AI REGIO platform has been developed, It is necessary to manage user access to the experiment application, ensuring data privacy and security.*
CL.01.EXP.15.IN.02	*Data collection and preparation*	*Exp05 is sharing the same requirements than those listed in cluster 3 table*
CL.01.EXP.02.IN.01	*AI technologies*	*The experiment is requiring:* *- Speech-to-Text and Text-To-Speech algorithms.* *-Content/Semantic Recognition* *-Rules based systems* *-Self-learning systems*
CL.01.EXP.03.IN.01	*AI technologies*	*The experiment is requiring:* - *Maze generation algorithms for automating the generation of 3D routing* - *Rule based systems*
CL.01.EXP.04.IN.01	*AI technologies*	*The experiment is requiring:* *-Machine learning: supervised learning algorithms* *-Expert systems: decision support systems* *-Recommendation engines* *-Natural Language Processing (NLP): Digital Assistants Chatbots, Content Generation*

Cluster 2: Factory Efficient and Sustainable Manufacturing

Table 2 - Requirements for AI REGIO Platform based on Cluster 2 needs

Identifier	High Level Requirement	Description
CL.02.EXP.05.IN.01	*Collect and prepare historical information*	*(Data4AI/AI Data Spaces for Manufacturing): It is necessary to store all the information related to the historical production orders that meet a common DataQuality criterion to ensure that errors are not going to be fed into the system and that all the stored information is compatible with future insertions of data.*
CL.02.EXP.05.IN.02	*Experiment and develop multiple planning and scheduling solutions*	*(AI4Manufactory/DIHIWARE): Experimentation with multiple AI techniques must help and encourage finding the best possible solution for planning and scheduling module for production process.*
CL.02.EXP.05.IN.03	*Domain knowledgeable agent*	*(Collaborative Intelligence Platform): The entire system must be monitored by the human interaction meeting a human-centred workflow in which the human interventions end up helping the AI system to suggest solutions in similar scenarios.*
CL.02.EXP.10.IN.04	*Collection of congruent heat demand data*	*(Data4AI/AI Data Spaces for Manufacturing): Historical data from various district heating systems must be collected and grouped under a single DataQuality criteria to ensure information coming from different sources can be used to extrapolate patterns and behaviours.*
CL.02.EXP.10.IN.05	*Real-time heat demand continuous prediction & model update*	*(AI4Manufactory/DIHIWARE): Multiple demand-based AI techniques must be used to end up developing a solution capable of providing a visual representation of the continuous prediction of the heat demand for a certain time range in the future.*
CL.02.EXP.10.IN.06	*Heat demand visualization and control*	*(Collaborative Intelligence Platform): The information provided by the AI system must be displayed in a visualization aid tool to visualize the AI system as well as to define the control settings.*
CL.02.EXP.12.IN.07	*Extraction and collection of correct interactive behaviour*	*(Data4AI/AI Data Spaces for Manufacturing): A common DataQuality which unifies information from the real factory shop floor and digital twin representations. Thanks to this, all the information and inconsistencies found in the real factory will be easily compared and updated in the digital twin representation.*
CL.02.EXP.12.IN.08	*Reinforcement learning mechanism definition*	*(AI4Manufactory): Reinforcement Learning techniques will be used to reach a simulation model capable of detecting non-rewarding actions.*
CL.02.EXP.12.IN.09	*System's AI model interaction with the user*	*(Collaborative Intelligence Platform): The resulting solution needs to be capable of being triggered and monitored by a user based on a set of production orders.*
CL.02.EXP.17.IN.10	*Extraction of tooling's performance in the production line*	*(AI Data Spaces for Manufacturing): The information relative to tooling's lifetime, performance and replacement must be accessible through an IDS solution. With this premise, the data is ensured to be*

		managed by its owner meanwhile AI solutions can be developed with it.
CL.02.EXP.17.IN.11	*Production line optimization based on costs and quality*	*(AI4Manufactory/DIHIWARE): Multiple AI techniques must be experimented to reach an AI based optimization tool which must be capable of ensuring the production line tooling's costs is reduced.*
CL.02.EXP.17.IN.12	*Production line optimization representation*	*(Collaborative Intelligence Platform): Thanks to this information, the system must be able to generate a warning in a monitoring system that helps to resolve the situation before the machinery of the production line is damaged.*

Cluster 3: Quality Control and Predictive Maintenance

Table 3 - Requirements for AI REGIO Platform based on Cluster 3 needs

Identifier	High Level Requirement	Description
CL.03.01	*Collect and store past data*	*(Data4AI/AI Data Spaces for Manufacturing): It is necessary to store all past data to enable creation of useful knowledge from it*
CL.03.02	*Support data acquisition from various data sources*	*(Data4AI/AI Data Spaces for Manufacturing): It is necessary to enable connection to and gathering of data from various sources (multimodal)*
CL.03.03	*Ensure Data Quality*	*(Data4AI/ AI Data Spaces for Manufacturing): It is necessary to ensure that the data used in experiments is of proper quality*
CL.03.04	*Flexible Data Pre-processing*	*(Data4AI): It is necessary to ensure that the data will be prepared for AI tasks in a flexible way (according to the requirements for AI)*
CL.03.05	*Orchestrated data preparation pipelines*	*(Data4AI): It is necessary to enable an easy composition of the data preparation pipelines*
CL.03.06	*Monitoring Human-Machine Interaction*	*(Collaborative Intelligence Platform): It is necessary to monitor human-machine interaction in order to enable its continual improvement*
CL.03.07	*Enabling scenarios with less data (AI-based)*	*(Data4AI/AI4Manufacturing): It is necessary to enable learning from scarce datasets*
CL.03.08	*Enabling complex orchestration of AI methods*	*(AI4Manufacturing): It is necessary to enable the combinations of different AI methods to provide the required processing (problem-solving)*

Cluster 4: Robotics and Human interaction

Table 4 - Requirements for AI REGIO Platform based on Cluster 4 needs

Identifier	High Level Requirement	Description
CL.04.EXP.01.IN.01	*Improve Data Quality*	*(Data4AI/AI Data Spaces for Manufacturing): Given some difficulties in labelling the metal sheets in the warehouse, nowadays, the data quality of the sheets (and other parts) in the warehouse can be poor. For this reason, today a lot of time is spent finding the right sheet within the piles.*
CL.04.EXP.01.IN.02	*Identifying the unique sheet to support operators*	*(AI4Manufactory/DIHIWARE): Implementation of a AI system able to support operators in the warehouse to identify exactly the unique sheet they need at that time for further delivery to production line.*
CL.04.EXP.01.IN.03	*Support humans to select sheets*	*(Collaborative Intelligence Platform): The AI system is expected to assist humans in the process of identification and selection of the needed sheets in the warehouse.*
CL.04.EXP.6.IN.04	*Data quality check*	*(Data4AI/AI Data Spaces for Manufacturing): To ensure the data quality, the sensor should be synchronised, and some data pre-processing should be implemented to deal with the noise and dirty data. Data quality influences the representation of the system properties and is the basic of the AI based experiment*
CL.04.EXP.6.IN.05	*AI-trajectory optimization*	*(AI4Manufactory/DIHIWARE):* An *AI-based model can be trained to learn the best robot pose combinations during the entire trajectory due to different robot stiffness of different robot poses.*
CL.04.EXP.6.IN.06	-	*(Collaborative Intelligence Platform): The focus of the experiment is not the human-machine collaboration since the focus is on using AI-based method to optimize the robot milling quality.*
CL.04.EXP.8.IN.07	*Access and usage control of data produced by AR system*	*(Data4AI/AI Data Spaces for Manufacturing): The access and usage of data produced by AR system is managed by IDS.*
CL.04.EXP.8.IN.08	*Objective data acquisition from dynamic effects*	*(AI4Manufactory): The moonshot of this experiment is to work towards objective data acquisition from dynamic effects in the manufacturing process, for instance learning curve, fatigue, and operator feedback.*
CL.04.EXP.8.IN.09	*AR operator support system*	*(Collaborative Intelligence Platform): The AR operator support system will be helping operators in giving the right amount of suggestions to complete complex modules.*
CL.04.EXP.11.IN.10	Extraction of Matchmaking system results	*(AI Data Spaces for Manufacturing):* The results of the Matchmaking system are exposed through an IDS connector
CL.04.EXP.11.IN.11	Matchmaking System and Web Service	*(AI4Manufactory/DIHIWARE): The Matchmaking system finds and creates manufacturing resource combinations, which can do the required process step(s) for a given product requirement.*

CL.04.EXP.11.IN.12	*Search and filtering of the feasible resources*	*(Collaborative Intelligence Platform): The system does the time consuming and cumbersome search and filtering of the feasible resources from large search spaces.*

3.1.2 Artificial Intelligence Requirements

Integrating Artificial Intelligent[5] (AI) assets into manufacturing environment is a complex task that requires a solid understanding of how these different domains complement and interact with each other. Currently in a general AI data-driven pipeline (Figure 13), the AI solutions are developed with known historical data and, once deployed, AI solutions continue operating with updated data. Moreover:

- **Data Ingestion**: Injecting the data from various sources into the system (considered as raw data)
- **Data Transformation**: Using statistical data profile and domain knowledge for data cleaning.
- **Past Data**: keeps a huge amount of old data in an efficient way that at the same time minimizes the amount of data and maximize the performance of the models.
- **Data Exploration**: Using various techniques (Normalization, Conversion like Wavelet, Dimension reduction) to prepare data in a way specific to analysis and training of the digital models.
- **Model Training**: Complex and time/computation intensive process of creating models/knowledge out of past data.
- **Real-World Usage**: it's the problem-solving task that affects the real-world operations of the generated models. It is usually realized as an interaction between the real time data and the learned models driven by AI.

In summary, there are two key dimensions in the pipeline:

- *ON line upper area*, related to the runtime operation, reflecting the real-world behaviour of the AI-based solution in the manufacturing domain.
- *OFF line lower area*, related to the design-time (or batch processing), reflecting the behaviour of the system mainly in the digital world. Firstly, at design-time with "Initial RAW Data" and later on, batch processing with "New RAW Data" concerning update of model.

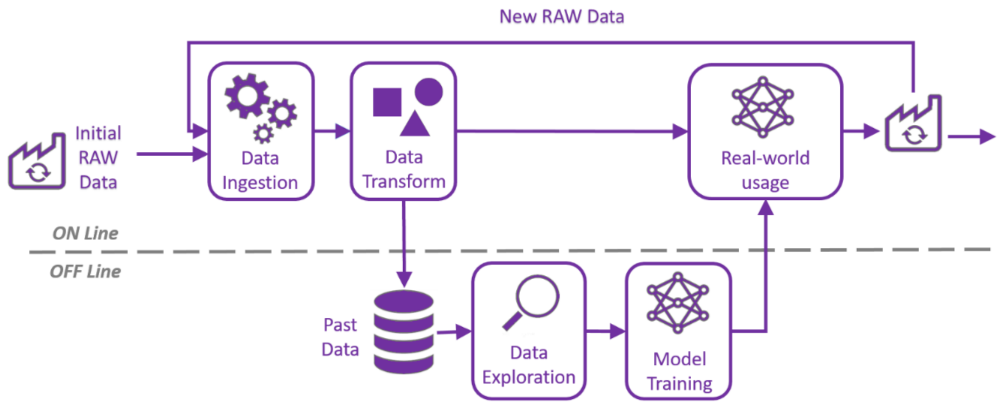


Figure 13 - AI data-driven pipeline

[5] As defined by High-Level Expert Group on AI: Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions.

With this scenario in mind the challenge then is establish the AI's high level "needs" as background to work properly with AI assets. In that sense, Data and AI landscape 2020[6] show us some Open-Source technologies that are in the scope of AI REGIO (Figure 14). All these Open-source existing solutions must be analysed to better design AI REGIO platforms, following a 'not building from scratch' strategy.

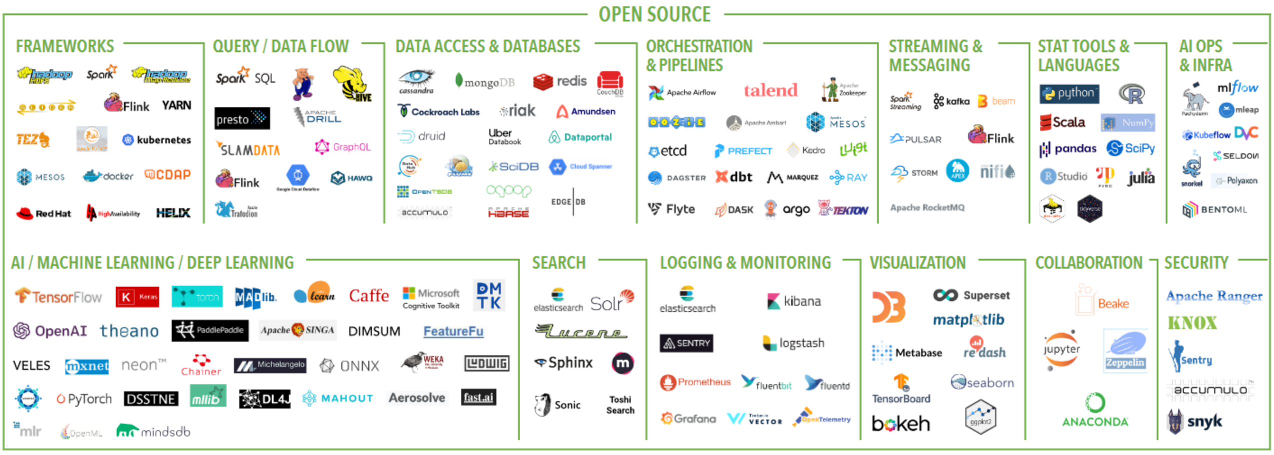


Figure 14 - AI Technologies Landscape 2020

In summary, the AI high-level requirements can be summarized as follows:

Table 5 - AI Requirements for AI REGIO Platform

Identifier	High Level Requirement	Description
AI.01	*Collect appropriate data*	*Machine Learning (ML) and Deep Learning (DL) require appropriate amounts of training data to avoid overfitting/underfitting of the AI models, while the overall effectiveness of AI depends on the diversity and dynamicity of input features.*
AI.02	*Efficiently query and prepare high quality data*	*Efficient exploration of available data through search engines, data querying and slicing mechanisms, with the ability for increasing data quality, and performing simple and complex operations (e.g. joining multiple datasets) with low-response times is a prerequisite for the execution of AI tasks.*
AI.03	*Support different AI libraries and frameworks*	*ML or DL models that provide solutions for a multitude of use cases and manufacturing problems, need to be typically built with popular frameworks (e.g. Keras, Tensorflow) and libraries (ex. Sk-learn), supporting high customizability and flexibility. Each model should be accompanied by the relevant evaluation metrics that will allow users and data analysts assess its effectiveness in terms of accuracy and decide on its adoption or even its improvement.*
AI.04	*Design and experiment with AI pipelines to address concrete manufacturing problems*	*ML/DL models are not applied in an isolated manner but need to be part of broader operations – and chains of operations – that involve data transformation and exploration and model training, apart from the model application and evaluation. Such pipelines are essentially different workflows designed to address different use cases and problems that bring their own datasets, to further promote re-use while capturing the*

[6] https://mattturck.com/wp-content/uploads/2020/09/2020-Data-and-AI-Landscape-Matt-Turck-at-FirstMark-v1.pdf

Identifier	High Level Requirement	Description
		sequences and dependencies among the involved tasks. A set of relevant functionalities for the management, orchestration, deployment and access to the pipelines should be provided to support different testing and experimentation 'playgrounds'.
AI.05	*Execute AI pipelines in production environments*	*The execution of AI pipelines needs to be scheduled and automated in order to leverage the available computation resources in the production environment (e.g. manufacturer's private cloud or on-premise infrastructures) in an optimal manner. The AI pipelines and their tasks need to be properly containerised for scaling with the help of AIOps technologies while ensuring their interoperability and compatibility to the extent it is possible.*
AI.06	*Understand AI results through human-machine interfaces*	*The results of an AI pipeline need to bring added value to the target end-user in the manufacturing use cases, through the easy exploration and comprehension of the inputs and outputs from the execution of an AI workflow. A user needs to be able to select the desired output format of the analysis results, e.g., in the form of visualisations, KPIs or predictions, powered by relevant libraries that provide lightweight data visualisations in appropriate human-machine interfaces.*
AI.07	*Share data, trained AI models and pre-configured AI pipelines*	*Following the FAIR data principles, the provided datasets, but also trained ML/DL models and preconfigured AI pipelines can be shared with other data scientists and manufacturers but also among DIHs. Explicit information needs to be provided in this case about the performance and evaluation metrics of the AI models. Access and exploration of existing pipelines, the acquisition of existing pipelines and the uploading of the users' data to run a selected or created pipeline need to be ensured always respecting the specified IPRs by their respective providers.*
AI.08	*Integration with the AI4EU platform*	*At high level, AI REGIO "needs" to interact properly with AI4EU platform. Becoming familiar with ACUMOS technology is essential to interoperate with the AI4EU experimental section (where the user can interactively create and configure AI pipelines). A complete list of video tutorials on how to use ACUMOS in the AI4EU platform can be found here.*

3.1.3 AI REGIO Sub-Systems

This section introduces the background, the main requirements and a conceptual architectural description for each AI REGIO sub-system.

AI REGIO has 5 Sub-Systems, which are:

1) Data4AI Platform, which is a new generation of platforms for ensuring data quality for AI applications based on syntaxis and semantic context.

2) AI4Manufacturing toolkit, which is expected to be a set of data analytics tool and techniques where prepared and cleaned data can be exploited for enabling experimentation with AI technologies.

3) DIHIWARE platform, which is an environment where providers and consumers of digital technologies related to AI development and adoption cannot just matching assets and needs, but they can collaborate to boost innovation.

4) Collaborative Intelligence Platform, which allows to design, monitor, optimize, and simulate the orchestration of human-centred processes workflow in terms of process management and Human-AI interaction.

5) AI Data Spaces for Manufacturing, which is data sovereignty solution based on Industrial Data Space (IDS) open-source current initiatives.

3.1.3.1 Industry 5.0

3.1.3.1.1 Rationale and SoTA

AI REGIO aims to provide the modelling and simulation playground for Industry 5.0 and Human-AI Collaborative Intelligence workflows within a fit for purpose platform. The latter will allow to support Industry 5.0 oriented scenarios, set in contexts where AI-driven autonomous systems are efficiently and effectively interacting with Humans according to the Collaborative Intelligence paradigm. The developed models will be suitable to support orchestration of human-centred processes in terms of process management and Human-AI interaction assessment, enabling also comparison of different technological solutions and selection in accordance to users needs.

In particular, a conceptual model for Human-AI interaction in Industry 5.0 is conceived and designed, which is focused to orchestrate, monitor, and simulate the human-centred processes. The AI REGIO Industry 5.0 model is based on the Digital Twin (DT) which is becoming a consolidate technology to simulate the physical industrial asset performance, thus allowing to predict failures or investigate problems. Specifically, the DT represents a virtual and faithful mirror of the physical process that allows to monitor the process parameters, to compare them with any analytic models, and to supply in real-time specific variations of parameters to keep the process always in optimal conditions.

The AI REGIO Industry5.0 model can be adopted as a reference architecture for a platform that promotes harmonization and orchestration between machines and the human factors, especially considering the cognitive and physical workload related to manufacturing operations. In addition, an implementation adhering to the proposed conceptual model's specifications will be provided, and this implementation will also be validated within a real case study, thus allowing to demonstrate the correctness of the overall proposed approach.

3.1.3.1.2 AI REGIO Requirements

The list of requirements is given in the following table.

Table 6 - Industry 5.0 Requirements

ID	Type	Description	Source (GA, D2.x, Survey, …)	Priority	Stakeholders
FR.1	Functional	It is needed the implementation of an orchestrator of human-centred processes within a graphical environment, where the orchestrator plays the role of balancing between the human	D2.5	H	Process owner

ID	Type	Description	Source (GA, D2.x, Survey, …)	Priority	Stakeholders
		component and machine within a human-centred process.			
FR.2	Functional	The platform should allow to simulate different possible solutions of configuration of the Collaborative Intelligence interactions, by changing the percentage of work performed by humans and machines, in order to balance between high value added tasks (requiring human brainpower and creativity of humans) and repetitive tasks (requiring high speed, precision and security typical of the machines). In addition, it should be also possible to simulate the type of planned interactions between humans and machines in order to enable a proactive evolution of the configured collaboration.	D2.5	H	Process owner
FR.3	Functional	An efficiency checker of each designed human-centred process. This functionality allows to check the efficiency of each designed process, in order to promote harmonization between machines and the human factors, especially considering the cognitive and physical workload related to manufacturing operations. For this reason, a quantitative modelling of the processes used in order to simulate and to assess all the possible scenarios. Some KPIs that can be used are Flexibility, Speed, Scale, Decision-Making, Personalization.	D2.5	H	Process owner
FR.4	Functional	The platform should also enable the comparison of different technological solutions of configuration and selection in accordance to the requirements of a specific scenario.	D2.5	H	Process owner

ID	Type	Description	Source (GA, D2.x, Survey, …)	Priority	Stakeholders
FR.5	Functional	In order to support the pattern "Machines Assist Humans", the platform should endow each machine of a learning model, through which machines learn from interplay with humans.	D2.5	H	Process owner

3.1.3.1.3 Architecture Description

The impact of the Collaborative Intelligence concept on the traditional model of the DT is studied with the overall aim to make this model compliant with the Industry 5.0 paradigm. In this context, the study investigates which are the new blocks of the conceptual model and which are the interrelations that involve these blocks. Since these interrelations also imply new streams of factory telemetry data, the study continues with the analysis of these data flows and in particular with the analysis of a software infrastructure that enables pipelines for the telemetry data.

Specifically, the new proposed model of DT, which is reported inFigure 12, introduces the blocks corresponding to the physical workers and their corresponding Personal Twin (PT). In addition, the model introduces the different blocks interrelations deriving from the two interaction patterns of the Collaborative Intelligence. These interrelations (represented in figure as brown hatched arrows) are then also characterized on the basis of the type of interactions of the workers with the assets (physical, visual, voice, digital). Thanks to these interrelations, humans and machines work and interact collaboratively in order to create (physical) products. In addition, the new proposed model of the DT analyzes the data pipelines through which information flow between two different components of the model. In Figure 12, these pipelines are represented as blue solid uni and bidirectional arrows.

In order to to support the implementation of an Industry 5.0 platform, different architectural components are needed and are described in the following. In addition, all these components are also reported in Figure 13, which emphasizes the role of pillar for the Industry 5.0 platform played by these components.

(C1) Industry 5.0 data space. The DT concept requires a homogeneous perspective of the handled CI information persisted across the different internal and external boundaries. Indeed, this information should not only be available vertically between value chains, but also distributed horizontally in an **Industry 5.0 data space** across organizational boundaries, where different assets (machines, workers, etc.) and the various architectural components of the DT (including the legacy databases) can exchange data while maintaining the sovereignty over them. In order to realize this horizontal data space, **two further architectural components** are needed:

(C1.1) Industry 5.0 Semantic data model and Ontology. This component aims at enhancing the interoperability among assets involved in the CI workflows and acting both as data producers and consumers. In addition to the definition of the semantic model, in order to enable interoperability within and across organizations, DT can also provide a **standardized application programming interface (API)** to manage the data access and usage in a trusted ecosystem (C1.2).

(C1.2) Trusted and secure ecosystem. This ecosystem would enable the secure and standardized exchange and the easy linkage of data across different sectors.

(C2) Middleware for the real-virtual synchronization. The real-time monitoring of the current system capabilities enables to update the decisions about maintenance strategy, to predict failures, and to implement optimization strategies. In order to dynamically assess the performance of a real asset, it is essential to connect to the latter and then process its produced data in real-time. The DT of a physical asset is the faithful mirror of its real counterpart and can be exploited to its full potential only if the virtual and real asset are bidirectionally synchronized. To achieve this synchronization process, the technological infrastructure supporting the DT must have a **middleware** application as a pillar, thanks to which the data (so-called **Factory Telemetry**) have to be managed within a circular process in the two directions: from the real factory to the DT for feeding the Online DT and from the DT to apply actions to the real factory.

(C3) Time-series databases and other NoSQL databases. Another relevant challenge consists in the persistence and accumulation of the acquired data (historical data) for feeding the offline analysis. In this regard, cloud based systems have to be taken into account to ensure the horizontal scalability of storage, computation, and communication capabilities, and to decouple storage, data processing, and data management. It should be noted that Factory telemetry data is typically under the form of Time Series, i.e., couples of (timestamp, value). Such a data do not fit well with the well-defined and rigid structure of the more common databases (based on relational model). For this reason, the current efforts of researchers and technicians are addressed towards the implementation of so called **time-series databases** (e.g., InfluxDB, etc.), i.e. purpose-built databases to manage time-series data, which belong to the set of NoSQL database.

(C4) Multi-scale models simulating the interaction "Machines Assist Humans". It should be noted that the DT is represented by a model that must be precise and detailed in order to perform accurate simulations and evaluations. However, a detailed model synchronization process can involve significant processing cost, especially when coupled with a high sample rate. For this reason, it is important to identify the right compromise between the level of detail (granularity) of the DT and the efficiency of the synchronization process. One solution can be the development of **multi-scale models** for both live telemetry and historical data which **allow to simulate the two pattern interactions of the CI**. In particular, the channel that need to be simulated is the "Machines Assist Humans", while for the channel "Humans Assist Machines" the simulation is less interesting while it is needed a learning model for the DTA (C7). Specifically, for the interaction "Machines Assist Humans", it is useful to simulate and understand live (when the process is running) to what extent the machine has to assist the worker on the basis of the latter capabilities and other aspects such his level of fatigue. This simulation can be exploited to assess a physical interaction (as in robotics scenarios) or visual\digital interaction where behind a physical component there is the DT's model. The application of the simulation in the first case seems more important as it allows to evaluate the mechanical \ physical aspects, but anyway it interesting to explore the difference deriving from the application of the simulation in the two cases.

(C5) Data cleansing mechanisms. It is essential to study valid strategies to distribute the intelligence and data of the DT close to the various data sources, thus limiting the use of the transmission band. One potential solution is applying a **data cleansing mechanism** at the edge level and in particular on the data produced at the physical level, since the quality of the data generated by the sensors or machines is typically less good than the data produced by an AI module.

(C6) Orchestrator. It is needed the **implementation** of an **orchestrator** of human-centred processes, which plays the role of balancing between the human component and machine. In addition, the orchestrator will be paired with a workflow enactment service which allow to interpret a chain and control the instantiation of services and sequencing of activities.

(C7) Learning model for DTA to support the interaction "Humans Assist Machines". DTA must comprise a learning model, through which machines learn from interplay with humans.

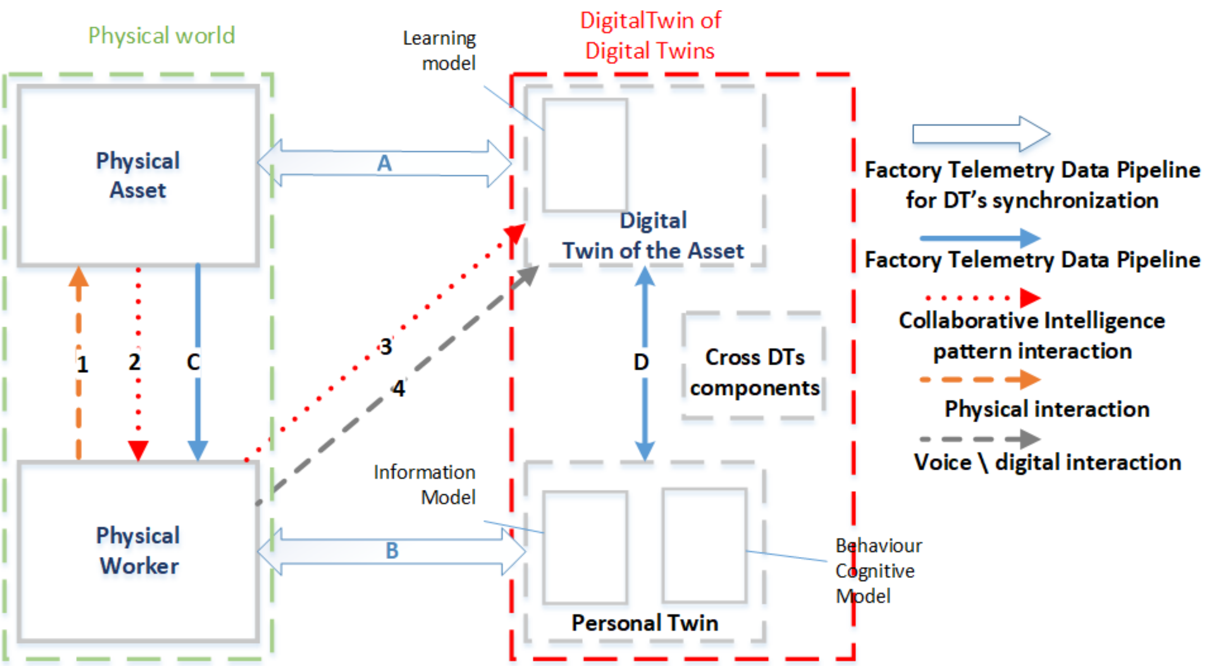


Figure 15 - Evolution of the DT model to become compliant with Industry 5.0 and Collaborative Intelligence

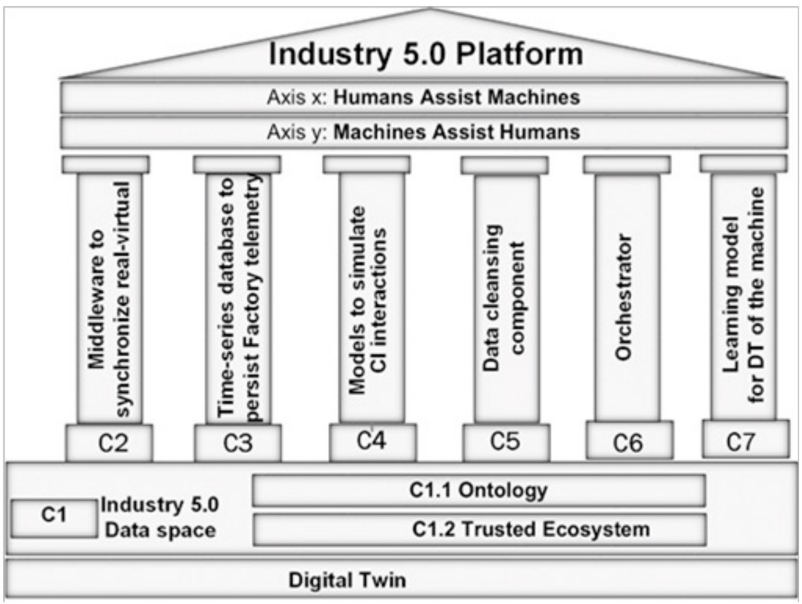


Figure 16 - Architectural components for the Industry 5.0 platform

3.1.3.2 Data4AI

3.1.3.2.1 Rationale and SoTA

Data4AI platform is a new generation of platforms for ensuring data quality for AI applications.

The motivation for Data4AI Platform is very clear:

- IBM: almost 80 percent of work involved in AI projects is data preparation
- PWC: while 76 percent of businesses aim to leverage their data to extract business value, only 15 percent have access to the appropriate type of data to reach that goal.
- University of Texas researchers have claimed that an increase of 10 percent in data usability will boost “annual revenue” by over $2 billion

That data collection and preparation is very challenging. Some of the issues are:

- problems with data consistency and quality,
- holes in coverage,
- undetected issues with instrumentation,
- faulty pre-processing.

Since these pitfalls are so common, it is needed to keep data collection on track, in terms of both scope and budget, and to help avoid these common pitfalls.

3.1.3.2.2 AI REGIO Requirements

The list of requirements is given in the following table:

Table 7 - Data4AI Requirements

ID	Type	Description	Source (GA, D2.x, Survey, …)	Priority	Stakeholders
R1.1	Functional	Enabling to upload a dataset	Analysis of the requirements of potential users	H	Process owner
R1.2	Functional	Enabling to define a set of criteria	Analysis of the requirements of potential users	H	Process owner
R1.3	Functional	Enabling to upload the information about the errors/anomalies which happened in the past	Interview with potential users	H	Process owner

ID	Type	Description	Source (GA, D2.x, Survey, …)	Priority	Stakeholders
R1.4	Functional	Providing visual presentation of results	Interview with potential users	H	Process owner
R2.1	Functional	Enabling to select the data source	Analysis of the requirements of potential users	H	Data Quality engineer
R2.2	Functional	To define a set of criteria for data quality	Analysis of the requirements of potential users	M	Data Quality engineer
R2.3	Functional	Enabling to check / validate created pipeline	Analysis of the requirements of potential users	H	Data Quality engineer
R2.4	Functional	Provide automatic reporting about the data quality	Interview with potential users	H	Data Quality engineer
R2.5	Functional	Provide report about the data preprocessing results	Interview with potential users	H	Data Quality engineer
R2.6	Functional	Enable reconfiguration of the pipeline	Analysis of the requirements of potential users	M	Data Quality engineer
R2.7	Functional	Enable automatic validation of the pipeline (based on defined KPIs)	Analysis of the requirements of potential users	M	Data Quality engineer

3.1.3.2.3 Architecture Description

High level conceptual model of the platform is presented in the following figure

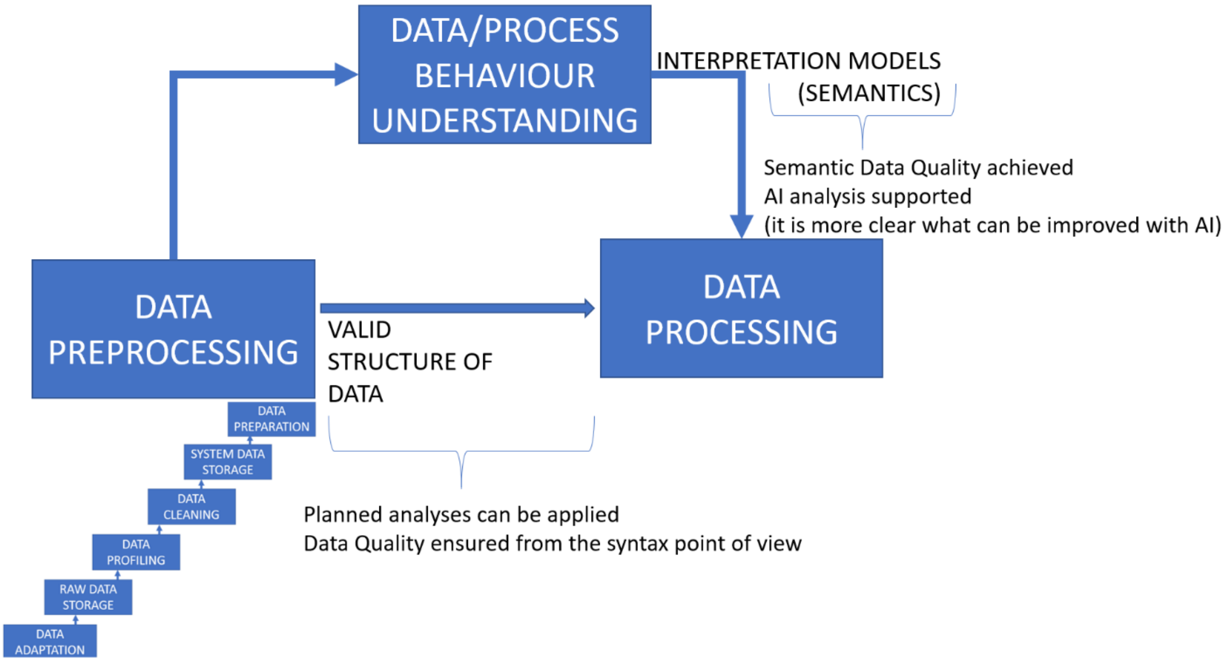


Figure 17 - Data4AI platform – high level conceptual model

The model is briefly described as:

- Data Preprocessing is a processing pipeline which transforms raw data in the well-formed data (valid structure) that can be processed by various data analysis methods
- Data/Process behavior understanding provides a view on the characteristics of the underlying process (the data is originating from)
- Data Processing is data analysis which can be done within or outside Data4AIPlatform

As depicted in the figure, Data Preprocessing ensures the Data Quality from the syntax point of view (it is in the valid form and can be processed automatically), whereas Data/Process behavior understanding provides a semantic context for data quality, i.e. for interpreting data in a way suitable for creating AI applications.

In the Figure 18 the Data Preprocessing pipeline is described.

In particular, describe the particular steps:

- Adapter – Reads the raw data from the files (properly prepared) and writes data into the raw data storage
- Raw data storage – Stores the raw data from the provided dataset into the previously defined format (d2msa, d2twin, etc.)
- Profiling – Data Inspection (overview profiling of the raw data stored in the raw data storage)
- Data cleaning – Data cleaning according to the info provided from data profiling (removing irrelevant data from the raw data)
- System data storage – Stores cleaned data after the profiling is done
- Data preparation – Getting data from the system data storage and preparing the data for the analytics algorithms

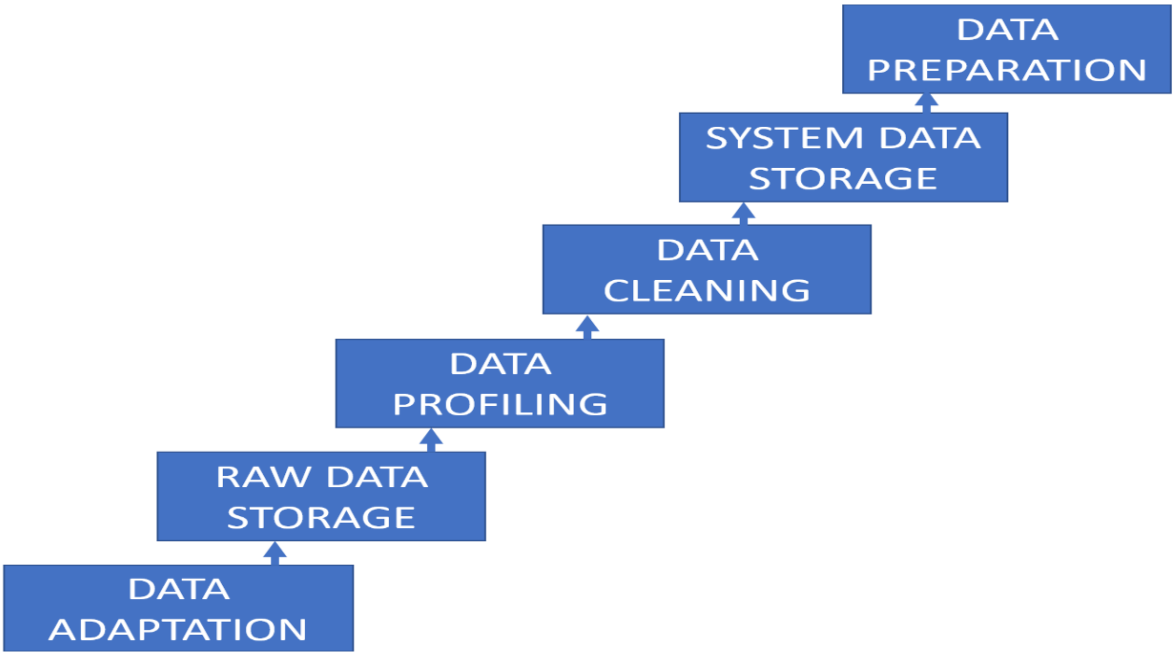


Figure 18 - Data Preprocessing pipeline

3.1.3.3 AI4ManufacturingToolkit

3.1.3.3.1 Rationale and SoTA

The AI4Manufacturing Toolkit (or platform) is expected to be a set of data analytics tool and techniques where prepared and cleaned data can be exploited for enabling experimentation with AI technologies.

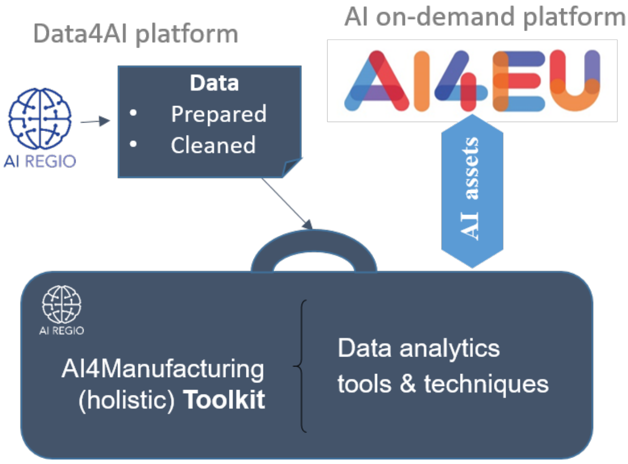


Figure 19 - AI4Manufacturing toolkit overview

AI4Manufacturing (holistic) toolkit can be seen as a set of technologies, tools and platforms. Moreover, it should be a collection of operational tools, designed to provide support to system integrators and technology adopters to create new AI-based applications. As a holistic toolkit including data analytics tools and AI techniques, AI4Manufacturing toolkit will provide services for boosting innovation in applying AI technologies in different business scenarios

As shown in Figure 19 two external components are key for the AI4Manufacturing toolkit, the Data4AI platform and, currently under development, the AI4EU platform:

- Data4AI platform (see section 3.1.3.2 for details) is a new generation of platforms for ensuring data quality for AI applications. AI4manufacturing toolkit should be able to use the prepared and cleaned dataset provided by Data4AI.
- AI4EU platform (www.ai4europe.eu) is a one-stop-shop for anyone looking for AI knowledge, technology, tools, services and experts. AI4Manufacturing toolkit should be interoperable by design and compatible with AI4EU platform. Figure 4 shows a simplified view of AI4EU platform where two main components have been identified:
    1. **AI4EU resource catalogues**, distributed in two locations:
        a. First one (based on Drupal) where user can find datasets, executables, AI models, docker containers, jupyter notebooks, as a service and libraries. For each asset, next information is available: description, documentation, performances, compliance with GDPR & ethics, IPR and support offered.
        b. Second one (based in Acumos) where AI dockerized models (as a docker container) and AI pipelines will be available. This catalogue contains information on the licensing and execution requirements of both reusable AI models and fully integrated solutions

       For both locations, an established, reviewed, publication process to guarantee the quality and the completeness of provided AI resources have been established.
    2. **AI4EU designer/experiment platform** (based on Acumos), where users can create on-board dockerized microservices and do visual composition of AI pipelines, based on Docker containers and gPRC as communications protocols. It is a complete system to package, share, license and deploy AI models into a full end-to-end solution.

This AI4EU designer/experiment section is not an execution environment and its goal is limited to the configuration/on boarding of services. To run the AI pipelines the user needs an external orchestrator as kubernetes or minikube.

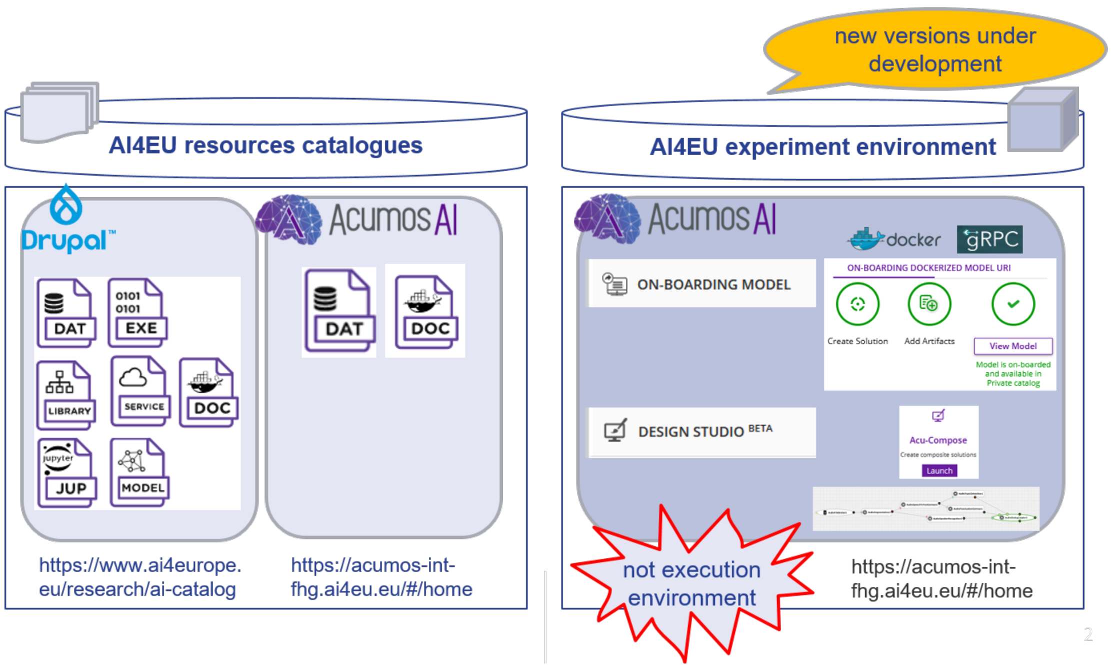


Figure 20 - AI4EU simplified view[7]

[7] AI4EU platform is a work in progress no new versions will be developed until the end of AI REGIO project.

In summary, the AI4Manufacturing toolkit intends to build the environment to enable the creation and experimentation of new AI-based pipelines with business sense. Its aim is to allow technical users to develop AI-based pipelines (from available tools) to generate new Data Analytics (Descriptive, Diagnostic, Predictive, Prescriptive, Machine Learning and AI, Deep learning, Statistics) solutions.

3.1.3.3.2 AI REGIO Requirements

Following list of requirements has been defined within AI4Manufacturing toolkit. There are 3 requirements related to management issues, 1 related to usability, 4 related to interoperability and 7 functional requirements:

Table 8 - AI4Manufacturing Toolkit Requirements

ID	Type	Description	Source (GA, D2.x, Survey, …)	Priority	Stakeholders
RM.01	Management requirement	As far as possible the platform must be implemented on top of existing Open-Source Software (OOS) technologies	D2.3	M	AI system developers/integrators
RM.02	Management requirement	The platform will be made of tools ("AI resources") implemented in different programming languages, as well as having different term of use, authorship and exploitation conditions.	WP4 meetings	H	AI system developers/integrators & AI REGIO management team
RM.03	Management requirement	The platform must follow a decentralized philosophy (as DIH philosophy), avoiding duplicities	WP4 meetings	M	AI REGIO management team
RU.01	Usability requirement	The user interface must be as simple as possible, to make easier the interaction with the platform	GA, D2.3	H	AI system developers/integrators
RI.01	Interoperability requirement	The platform must be able to use data and prepared data coming from	GA	H	AI REGIO management team

ID	Type	Description	Source (GA, D2.x, Survey, …)	Priority	Stakeholders
		Data4AI platform and other well-known data spaces			
RI.02	Interoperability requirement	The platform must be interoperable (if possible) with experimental AI4EU platform (based on ACUMOS) and other well-known AI pipelines designer tools.	GA, WP4 meetings	H	AI REGIO management team
RI.03	Interoperability requirement	The platform should be fully functional even without the connection with external platforms	WP4 meetings	H	AI REGIO management team
RI.04	Interoperability requirement	The platform should be interconnected with the AI REGIO portal (as well as with its "Technological Assets Catalogue" component)	WP4 meetings	H	AI REGIO management team
RF.01	Functional requirement	The platform must include an orchestrator of well-known AI pipelines designer environments, complementing them (avoid reinventing the wheel)	WP4 meetings	H	AI system developers/integrators
RF.02	Functional requirement	The platform must provide support for multiple workflow managers (freedom of choice), maximising reuse of backgrounds and existing components	WP4 meetings	H	AI system developers/integrators & AI REGIO management team
RF.03	Functional requirement	Each pipeline must be modelled as a direct acyclic graph, to be executed in a specific production environment, compatible with the designer	WP4 meetings	H	AI system developers/integrators & AI REGIO management team

ID	Type	Description	Source (GA, D2.x, Survey, …)	Priority	Stakeholders
		environment where they were created			
RF.04	Functional requirement	The output of an AI Pipeline can be specific results that are in the form of predictions, KPIs or required visualizations. Each AI Pipeline needs to have its own IPR (considering the IPR of all models included in the pipeline and the involved input/training data)	WP4 meetings	H	AI system developers/integrators & AI REGIO management team
RF.05	Functional requirement	The platform must have a pre-defined AI REGIO templates (e.g. predictive maintenance, defect detection, ...) to build AI applications within manufacturing domain	WP4 meetings	M	AI system developers/integrators & AI REGIO management team
RF.06	Functional requirement	The platform must guide the user in the selection of existing well-stablish technologies within open-source AI pipelines designer	WP4 meetings	M	AI system developers/integrators & AI REGIO management team
RF.07	Functional requirement	Only registered users are able to use the platform	WP4.3 leader's proposal	L	AI REGIO management team

3.1.3.3.3 Architecture Description

AI4Manufacturing platform must be a holistic toolkit focused on manufacturing and build on the top of well-known AI open-source solutions. Moreover, the platform should be, at least, interoperable with dataset coming from Data4AI platform, AI resources coming from IoT Catalogue, and with AI4EU resources. Following these principles, two main blocks within AI4Manufacturing architecture can be defined (Figure 3):

- An AI pipelines **orchestrator**, to provide support for multiple workflow managers, maximising reuse of backgrounds and existing components as well as providing more flexibility in developing new AI pipelines in your preferred technological stack.

- An AI pipelines **designer environment**, to ease the composition of AI resources in a workflow, maximising reuse of backgrounds and existing components as well as providing more flexibility in developing new AI pipelines in your preferred technological stack.
- In both cases one or more open-source technologies will be selected by AI REGIO technological partners to support the AI4Manufacturing platform.

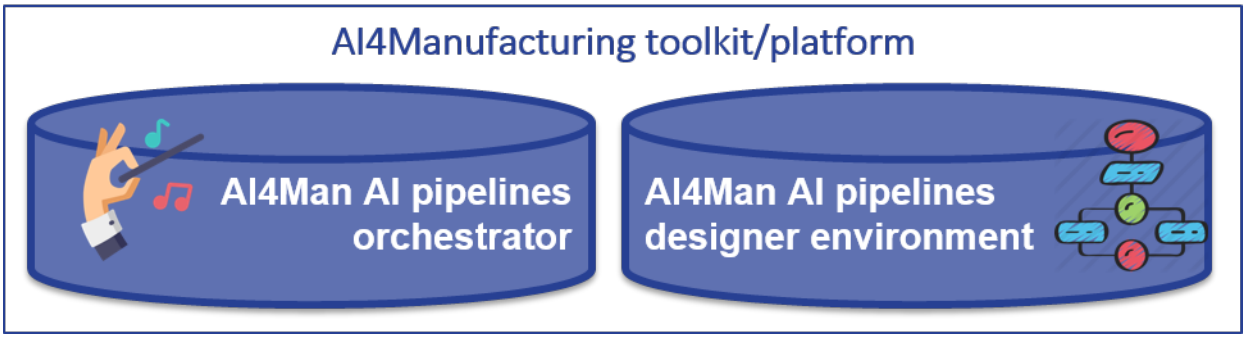


Figure 21 - High level AI4Manufacturing toolkit/platform architecture

3.1.3.4 DIHIWARE

3.1.3.4.1 Rationale and SoTA

The DIHIWARE Platform is a solution developed by ENG within the MIDIH H2020 EU project (http://midih.eu/) and currently in use in many ecosystems in Europe. The DIHIWARE Platform offers a complete collaboration environment inspired by Enterprise Social Software, realizing a bridge among stakeholders with different experiences backgrounds, providing access to the latest knowledge and expertise and pulling teams together and supplying a fertile ground for experimentation. These knowledge-driven services are fully integrated with collaborative services in order to create a digital space where all the AI REGIO stakeholders can collaborate to boost innovation together, enabling and digital supporting the co-creation methodology envisaged.

The main users of the DIHIWARE Platform can be reconciled to the following three families:

- Manufacturing SMEs (demand)
- Technology SMEs (offer)
- DIH (broker)

3.1.3.4.2 AI REGIO Requirements

Table 9 - DIHIWARE Requirements

ID	Type	Description	Source (GA, D2.x, Survey, …)	Priority	Stakeholders
FR-001	Functional requirement	The system must provide functionalities to	GA, D2.3	H	Technology SMEs (solution providers) and DIHs

ID	Type	Description	Source (GA, D2.x, Survey, …)	Priority	Stakeholders
		allow the federation of services and assets catalogues: create a common Knowledge base, a unique hub able to give information and contextualize entities.			
FR-002	Functional requirement	Advanced search options able to narrow the scope of user search query	GA, D2.3	H	Technology SMEs (solution providers), DIHs and Manufacturing SMEs (manufacturing Industries)
FR-001	Functional requirement	Application Interoperability: defines the tools, technologies and procedures associated with facilitating the interchange of information in different systems environments (Among DIHIWARE istances and with external systems e.g Iot Catalogue, AI4EU)	GA, D2.3	M	Technology SMEs (solution providers), DIHs and Manufacturing SMEs (manufacturing Industries)
C&C-001	Concept and Content requirement	User Segmentation: to divide potential users into meaningful subgroups based on their characteristics and determine access to content	GA, D2.3	H	Technology SMEs (solution providers), DIHs and Manufacturing SMEs (manufacturing Industries)

ID	Type	Description	Source (GA, D2.x, Survey, ...)	Priority	Stakeholders
C&C - 002	Concept and Content requirement	Matchmaking configuration: sets out rules for matching stakeholders and forming teams	GA, D2.3	H	Technology SMEs (solution providers), DIHs and Manufacturing SMEs (manufacturing Industries)
C&C - 003	Concept and Content requirement	Create custom-tailored environments where providers and consumers of digital technologies are not just matching assets and needs - according to the demand/supply of solutions, applications and technologies - but they are collaborating to boost innovations	GA, D2.3	H	Technology SMEs (solution providers), DIHs and Manufacturing SMEs (manufacturing Industries)
UX-001	Usability requirement	Make easier the interaction with the system revising the User Interface Design	GA, D2.3	H	Technology SMEs (solution providers), DIHs and Manufacturing SMEs (manufacturing Industries)

3.1.3.4.3 Architecture Description

The DIHIWARE Platform is an integrated environment made of two main systems: The collaboration portal and the Catalogues Management System. Each system provides a specific function and complements the functionality of other two systems. The Platform is also adopting state of the art identity management and authorization components in order to ease interoperability with other platforms using well known standards.

The Collaboration Portal will be the main entry point for users and enables open collaboration, online community building and management and access to knowledge. The Catalogue Management System provides functionalities related to both the back-end management, structure and storage of the catalogues as well as for the front-end (integrated within the KMS environment) allowing users to interact with them (e.g. view, filter and select).

Innovation Action - This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N. 952003

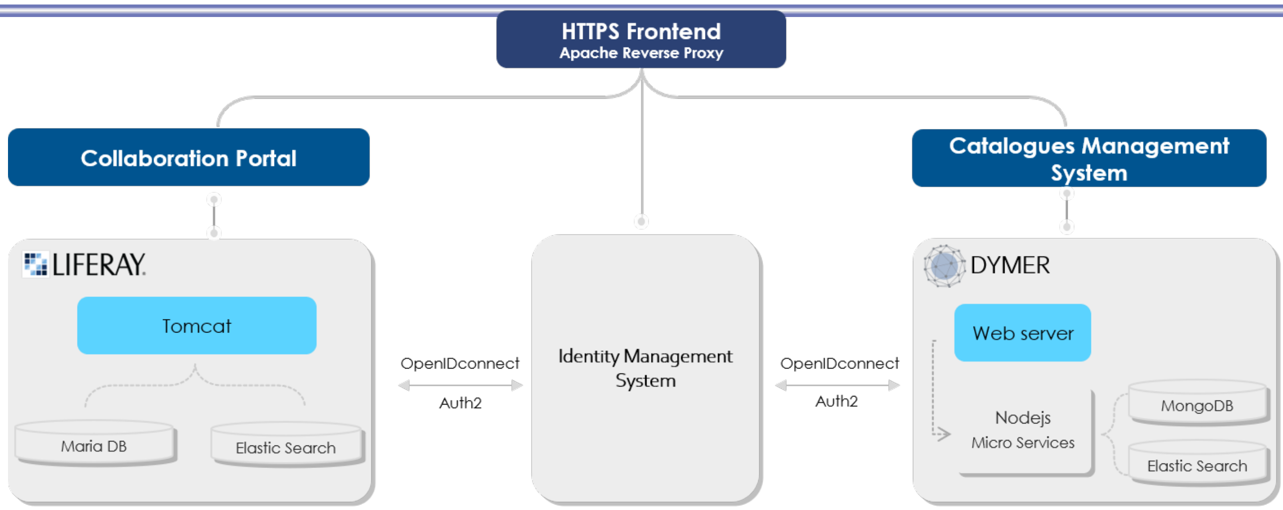


Figure 22 - High level decomposition DIHIWARE

The use of Open-Source components together with a flexible and modular integration and deployment approach guarantee the possibility to have custom-tailored solutions suitable for the variegated environments.

The high-level decomposition is shown in Figure 22 and each system (including its core technologies and functionalities) is described in more details in the following sub-sections.

3.1.3.5 AI Data Space for Manufacturing

The European strategy for Data makes a call for common rules and mechanisms that ensure that "data can flow within the EU and across sectors". The International Data Space (IDS) initiative identify three further design principles that take up on the demands for data sovereignty, data traceability and trust between the data space participants (Otto, 2020)

Three major functional aspects should be taken into account: Trust, Interoperability and Governance.

	Domain specific design principles for data spaces		
"How to realize data spaces"	Unlimited **Interoperability**	**Trust** between different security domains	**Governance** for the data economy
Domain specific components (building blocks)	Data standards and formats	Authorisation	Legal agreements
			Operational agreements
	Metadata	Certification and Monitoring	Earning models
	Data Exchange		Certification and monitoring
			Goveranance (e.g. AAS)

Figure 23 - Domain specific design principles for DS

3.1.3.5.1 Rationale and SoTA

Manufacturing sector is addressing several objectives:

- From mass production to mass customization.
- Reliable supply chains
- Servitization
- 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 type 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 horizontal and vertically.

The existing situation in the manufacturing sector is that, in 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.

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 lineal control and robotics

And manufacturing companies consider the impact of data in the following aspects:

- Cost reduction
- Quality improvement
- Servitization
- Supply chain management.

3.1.3.5.2 AI REGIO Requirements

Table 10 - AI Data Space for Manufacturing Requirements

ID	Type	Description	Source (GA, D2.x, Survey, …)	Priority	Stakeholders
FR-001	Functional Requirement	The system must provide a **broker service/module** in order to link consumers with providers together	GA, D2.3	H	Broker module provider at regional level and/or at international level. Manufacturing Industrial companies (data providers)

ID	Type	Description	Source (GA, D2.x, Survey, …)	Priority	Stakeholders
		with a catalogue of all services available			and app providers (data consumers) DIHs
FR-002	Functional Requirement	The system must provide an **execution core container (IDS connector)** in charge of the data exchange through the IDS ecosystem	GA, D2.3	H	IDS connector and ECC provider at regional level and/or at international level. Manufacturing Industrial companies (data providers) and app providers (data consumers) DIHs
FR-003	Functional Requirement	The system must provide a **clearing house service** in order to record all activities carried out in the data exchange course so that the transaction can be billed.	GA, D2.3	H	Clearing House module provider at regional level and/or at international level. Manufacturing Industrial companies (data providers) and app providers (data consumers) DIHs
FR-004	Functional Requirement	The system must support an **app store** for IDS data apps so that data transformation services can be deployed within the connector	GA, D2.3	H	App store provider (at regional or international level) App Providers
FR-005	Functional Requirement	The system must provide **usage control data application**	GA, D2.3	H	Manufacturing SMEs and Industrial companies as data consumers
FR-006	Functional Requirement	The system must provide a service to create, maintain, manage, monitor and validate the **identity** information of the participants in IDS in order to prevent unauthorized access to data	GA, D2.3	H	DAPS / IP module provider at regional level and/or at international level. Manufacturing Industrial companies (data providers) and app providers (data consumers) DIHs
FR-007		The system must provide a **vocabulary**	GA, D.23	H	

ID	Type	Description	Source (GA, D2.x, Survey, …)	Priority	Stakeholders
	Functional Requirement	**provider** that manages and offers vocabularies (i.e. ontologies, reference data models, or metadata elements) that can be used to annotate and describe datasets			Vocabulary provider module provider at regional level and/or at international level. Manufacturing Industrial companies (data providers) and app providers (data consumers) DIHs
FR-008	Functional Requirement	The system should provide a support to connect and retrieve data from AAS (Asset Administration Shell) components	GA, T5.4	M	Manufacturing SMEs, Manufacturing Industrial companies (data providers)

AI REGIO as a data space: Multiple customers (SMEs) or data providers, sharing their data with multiple Service providers (DIH or competence centres) through a network of European DIHs.

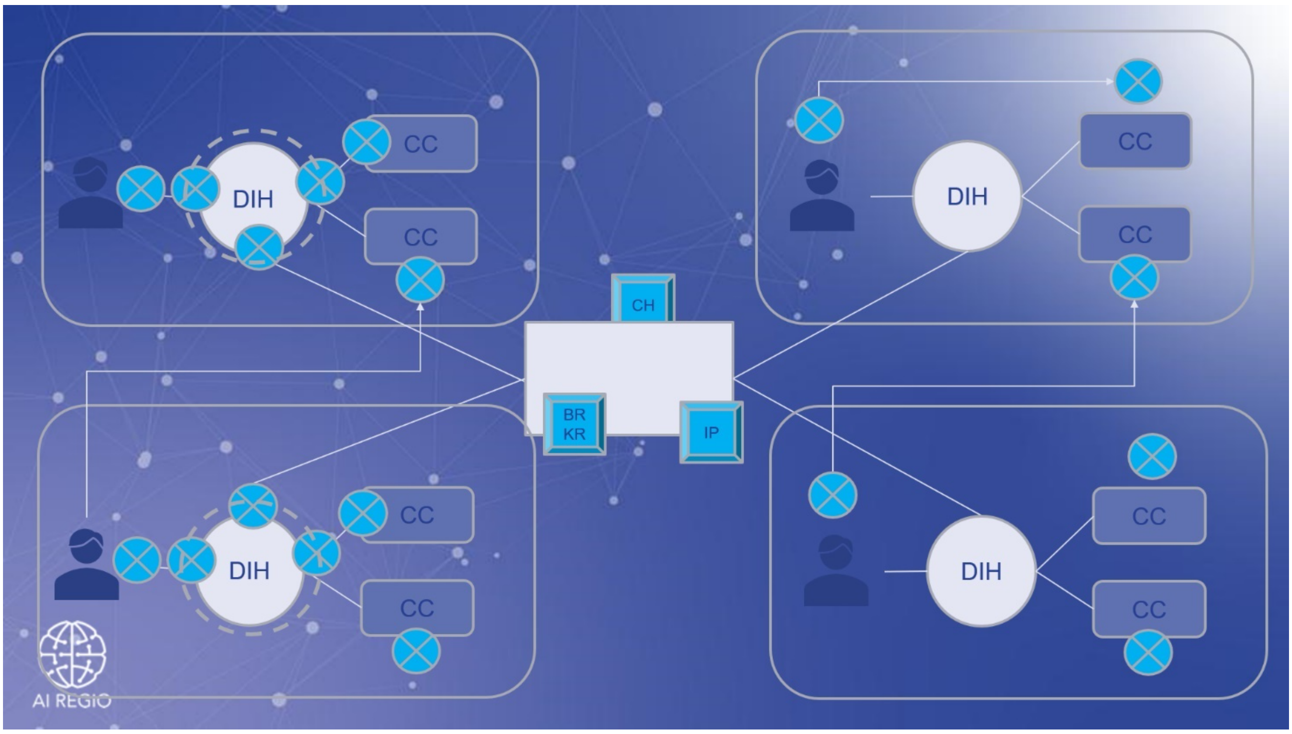


Figure 24 - AI data space for AI REGIO

3.1.3.5.3 Architecture Description

Architecture requirements for a data space:

- **Data sharing empowerment** is about ensuring that decisions can be made by appropriate stakeholders.
- **Data sharing trustworthiness** is about ensuring that data spaces operate according to expected requirements.
- **Data sharing publication** is about enabling data to be published so it can be easily located by data consumers.
- **Data sharing economy** is about creating the conditions for data trading.
- **Data sharing interoperability** is about providing the ability for all applications in data spaces to create, use, transfer and effectively exchange data.
- **Data space engineering flexibility** is about providing the ability for engineers to add customised features in data processing applications and data platforms.
- **Data space community** is about fostering maximum reuse of data space solutions.

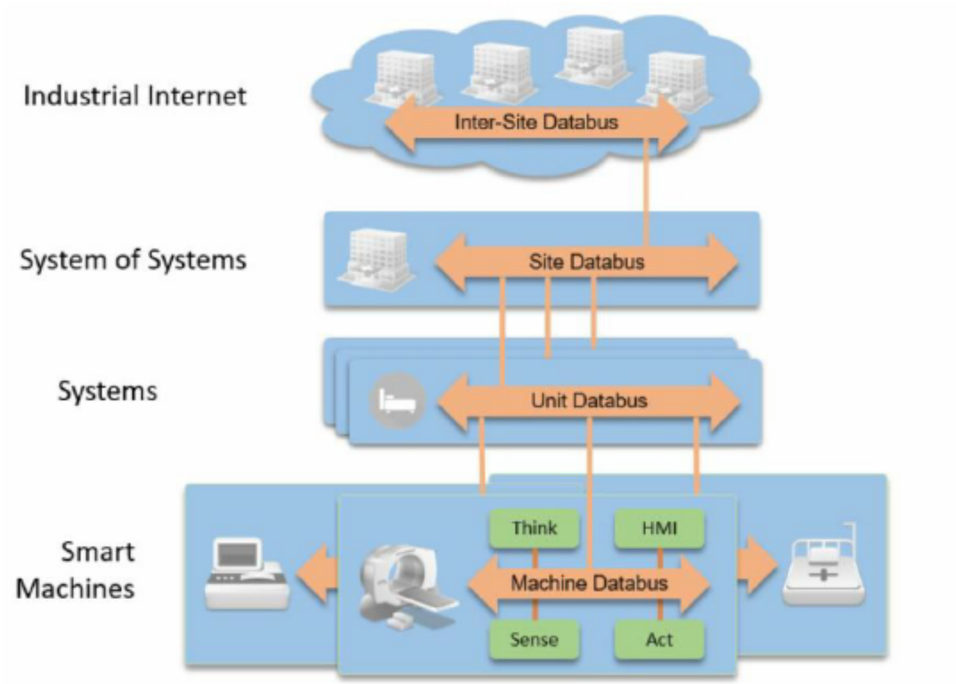
3.2 Design Principles for AI REGIO RA

The AI REGIO Reference Architecture will follow, like happen for the most of IIoT system implementations, the well-established architectural patterns described in relation with the benefits that can bring to the AI REGIO project:

- **Three-tier architecture:** It is an easy representation of an Industrial Internet solution, but simultaneously describes the main areas to be considered when design an Industrial Internet System. Three-tier architecture is composed by:
    - **Edge tier:** data from all the end nodes is collected, aggregated, and transmitted over the proximity network to a border gateway. Depending on the protocols and technologies used within the edge tier, some data translation and interface integration may be applied at hubs, remote I/O devices, or protocol convertors. The edge tier contains the functions for the control domain.
    - **Platform tier:** receives data from the edge tier over the access network and is responsible for data transformation and processing. The platform tier is also responsible for forward control commands from the enterprise tier to the edge tier. It is within the platform tier where it will be located the majority of the functions related to the information and operations domains.
    - **Enterprise tier:** implements domain-specific applications, applications and business logic for decision support systems and end-user interfaces to end-users including operation specialists. The enterprise tier receives data flows from the edge and platform tier. It also issues control commands to the platform tier and edge tier.

  They are connected by three networks (Proximity, Access and Service networks).

- **Layered Databus:** The Layered Databus architecture pattern is often applied to IIoT systems, as it provides low-latency what minimizes delays between the transmission and receipt of data, secure and peer-to-peer data communications across logical layers of the system. At the lowest level, the **Machine Databus** is used by smart objects such as: sensors, objects, devices, machines and products for local control, automation and real-time analytics. The Site and Unit layers could also be coincident and use another Databus for supervisory



control and monitoring.

Figure 25 - Layered Databus Architecture

Considering what is described above, a first high level of RA is designed, taking into account the core of the platform is composed by the 5 techno assets: Industry 5.0, Data4AI, AI4Manufacturing Toolkit, DIHIWARE, AI Data Space for Manufacturing.

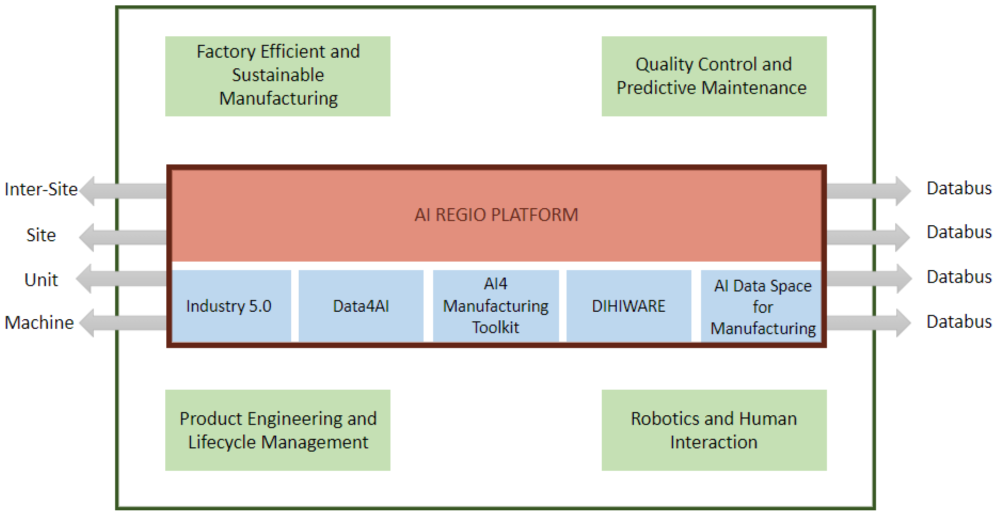


Figure 26 - AI REGIO High Level RA

A special focus must be done on how manage the data-in-motion and the data-at-rest to have a data-driven architectural approach, where:

- **Data in Motion:** platforms are software subsystems that perform real-time calculations on event data in motion. They retain a relatively small working set of stream data in memory for a limited time window, typically seconds to hours.
- **Data at Rest:** data, structured or unstructured, that has been flushed out from the memory and stored in any digital form (e.g. cloud storage, file hosting services, databases, data warehouses, …).

3.3 AI REGIO Reference Architecture

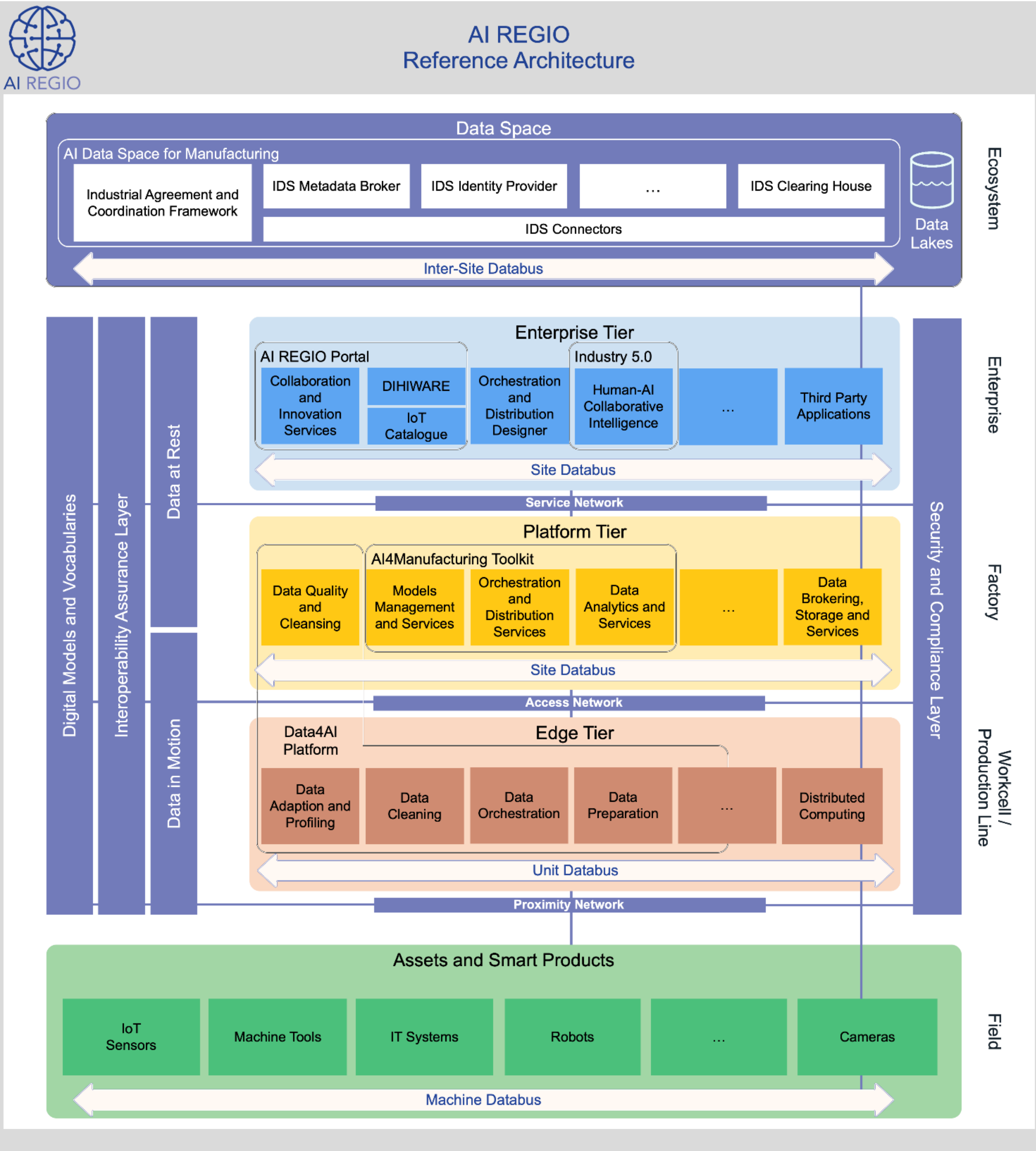


Figure 27 - AI REGIO Reference Architecture

- **Asset and Smart Products**: ecosystem of heterogeneous devices (IoT Sensors, Machine Tools, IT Systems, Robots, Cameras and so on) to gather machine data and make it available to the upper layers and IoT systems.

- **Edge Tier**: computes some data management and analysis functions, in small datasets, using data, applications, and services contained in the edge.
    - **Data adaption and profiling**: the service used for making the information available within its environment and adapt its execution accordingly.
    - **Data cleaning**: the process able to guarantee the correctness of a huge amount of data, starting from data mining.
    - **Data orchestration**: implements the logic for defining complex data flow processes between devices, data services, applications and people to produce desired outcomes.
    - **Data preparation**: in charge of data manipulation to prepare it in the format needed to be analyzed.
    - **Distributed computing**: is the link to the upper layer, allowing the data pipeline to communicate with the Platform services.

- **Platform Tier**: leverages the processing of data enables making smarter, and sometimes faster, business decisions.
    - **Data quality and cleaning**: the process to discover and repair corrupted or inaccurate data, ensuring the quality of information in the dataset.
    - **Model management and services**: the services used to manage the models that feeds the platform once the data are already processed and ready to be used.
    - **Orchestration and Distribution services**: services used to deploy the solution along the supply chain once it is designed thanks to the AI REGIO platform
    - **Data analytics and Services**: set of data analytics techniques and services used to take advantage from the information extracted/provided by the subsystems.
    - **Data brokering, storage and services**: interact with the lower layer to gather data and storage it in order to make it available for the services in the platform.

- **Enterprise Tier**: receives the data and integrates it with data from other systems, to perform analysis across business silos, carrying out industry domain-specific business applications, related decision support and business intelligence systems. This tier provides interfaces to human consumers.
    - **Collaboration and innovation services**: allow the connection and information exchange, changing the way in which is provisioned in order to increase the value of service offered.
    - **DIHIWARE**: an integrated solution aiming to provide access to the latest knowledge, expertise and technology during SMEs Digital Transformation pathways toward piloting, testing and experimenting new digital technologies.
    - **IoT Catalogue**: a catalogue of solutions already for which there is a description identifying the specific problems solved.
    - **Orchestration and Distribution Designer**: is the service used to design the solution along the supply chain.
    - **Human-AI Collaborative Intelligence**: referred to Industry 5.0, is the service offered to improve the interaction human-machine with the help of AI.
    - **Third Party Applications**: any other kind of application can integrate AI REGIO Platform.

- **Data Space**
    - **Industrial Agreement and Coordination Framework**: guidance to satisfy requirements in order to obtain a certification, including business, organizational and operational agreement.

- **IDS Metadata Broker**: is an intermediary that stores and manages information about the data sources available in the data space.
- **IDS Identity Provider**: offers a service to create, maintain, manage, monitor, and validate identity information of and for participants in the data space.
- **IDS Connectors**: is the technical component to standardize data exchange between participants in the data space.
- **IDS Clearing House**: is an intermediary that provides clearing and settlement services for all financial and data exchange transactions.

* **Digital Models and Vocabularies**: the set of models and vocabularies used in the among all the layers allowing the interoperability.

* **Interoperability Assurance Level**: ensures interoperability [10] among all the levels finding a link to connect them, allowing the exchange of information.

* **Data in Motion**: real time data, typically represents Industrial IoT data coming from sensors and devices.

* **Data at Rest**: historical data, typically stored in IT and legacy systems.

* **Security and compliance Level**: service offered to all the layer act to ensure a trusted environment at all the levels.

3.4 Reference Implementation

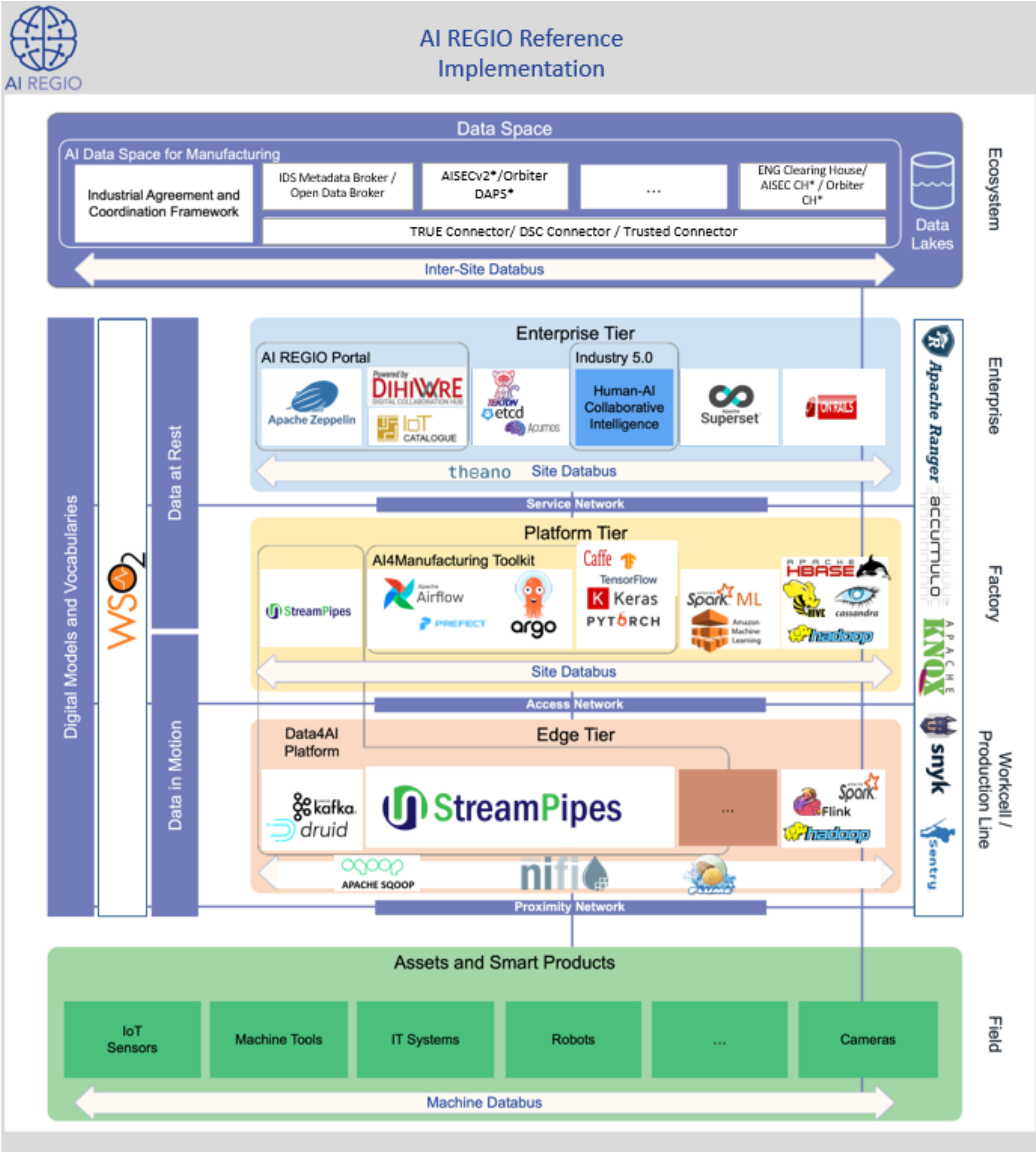


Figure 28 - AI REGIO Reference Implementation

- **SQOOP:** Scoop is a common ingestion tool that is used to import data into Hadoop from any RDBMS. Sqoop provides an extensible Java-based framework that can be used to develop new Sqoop drivers to be used for importing data into Hadoop. Sqoop runs on a MapReduce framework on Hadoop and can also be used to export data from Hadoop to relational databases.

- **Nifi:** Apache NIFI supports powerful and scalable directed graphs of data routing, transformation, and system mediation logic. Some of the high-level capabilities of Apache NiFi include Web-based user interface, Seamless experience between design, control, feedback, and monitoring, data Provenance, SSL, SSH, HTTPS, encrypted content, etc, pluggable role-based authentication/authorization.Apache nifi is highly configurable with loss tolerant vs guaranteed delivery, low latency vs high throughput, dynamic prioritization, flow can be modified at runtime, back pressure.
- **Flume:** A Java-based ingestion tool, Flume is used when input data streams-in faster than it can be consumed. Typically, Flume is used to ingest streaming data into HDFS or Kafka topics, where it can act as a Kafka producer. Multiple Flume agents can also be used collect data from multiple sources into a Flume collector.
- **Kafka:** Kafka is a distributed publish-subscribe messaging system which integrates applications/data streams. It is fast, scalable and reliable messaging system which is the key component in Hadoop technology stack for supporting real-time data analytics or monetization of Internet of Things (IoT) data. Kafka can handle many terabytes of data without incurring much at all. Apache Kafka is altogether different from the traditional messaging system. It is designed as a distributed system and which is very easy to scale out. Kafka is designed to deliver three main advantages over AMQP, JMS etc.
- **Druid:** Apache Druid is a database that is most often used for powering use cases where real-time ingest, fast query performance, and high uptime are important. As such, Druid is commonly used for powering GUIs of analytical applications, or as a backend for highly concurrent APIs that need fast aggregations. Druid works best with event-oriented data.
- **Streampipes:** StreamPipes is a self-service (Industrial) IoT toolbox to enable non-technical users to connect, analyze and explore IoT data streams. StreamPipes has an exchangeable runtime execution layer and executes pipelines using one of the provided wrappers, e.g., standalone or distributed in Apache Flink. Pipeline elements in StreamPipes can be installed at runtime - the built-in SDK allows to easily implement new pipeline elements according to your needs. Pipeline elements are standalone microservices that can run anywhere - centrally on your server, in a large-scale cluster or close at the edge.
- **Flink:** Flink is a streaming data flow engine which aims to provide facilities for distributed computation over streams of data. Treating batch processes as a special case of data streaming, Flink is effective both as a batch and real-time processing framework, but it puts streaming first. Flink offers several APIs which includes static data API like DataStream API, DataSet API for Java, Scala and Python and SQL-like query API for embedding in Java, Scala static API code. Flink also has its own machine learning library called FlinkML, its own SQL Query called MRQL as well as graph processing libraries.
- **Spark:** Apache Spark is an open-source parallel processing framework for running large-scale data analytics applications across clustered computers. It can handle both batch and real-time analytics and data processing workloads. It's part of a greater set of tools, including Apache Hadoop and other open-source resources for today's analytics community. In this way, it can be considered as a data analytics cluster computing tool. It can be used with the Hadoop Distributed File System (HDFS), which is a particular Hadoop component that facilitates complicated file handling.
- **Hadoop:** The Hadoop Distributed File System (HDFS) is a distributed file system designed to run on commodity hardware. It has many similarities with existing distributed file systems. However, the differences from other distributed file systems are significant. HDFS is highly fault-tolerant and is designed to be deployed on low-cost hardware. HDFS provides high throughput access to application data and is suitable for applications that have large data sets.

- **Airflow:** an open-source workflow management platform. Creating Airflow allowed Airbnb to programmatically author and schedule their workflows and monitor them via the built-in Airflow user interface. It is written in Python, and workflows are created via Python scripts. Airflow is designed under the principle of "configuration as code". While other "configuration as code" workflow platforms exist using markup languages like XML, using Python allows developers to import libraries and classes to help them create their workflows.
- **Argo:** Argoproj is a collection of tools for getting work done with Kubernetes like: Argo Workflows (Container-native Workflow Engine), Argo CD (Declarative GitOps Continuous Delivery), Argo Events (Event-based Dependency Manager), Argo Rollouts (Progressive Delivery with support for Canary and Blue Green deployment strategies).
- **Pytorch:** is an open-source machine learning library based on the Torch library (an open-source machine learning library, a scientific computing framework, and a script language based on the Lua programming language. It provides a wide range of algorithms for deep learning, and uses the scripting language LuaJIT, and an underlying C implementation), used for applications such as computer vision and natural language processing.
- **Keras**: is an open-source software library that provides a Python interface for artificial neural networks. Keras acts as an interface for the TensorFlow library. Designed to enable fast experimentation with deep neural networks, it focuses on being user-friendly, modular, and extensible. It was developed as part of the research effort of project ONEIROS (Open-ended Neuro-Electronic Intelligent Robot Operating System).
- **Tensorflow**: is a free and open-source software library for machine learning. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks. Tensorflow is a symbolic math library based on dataflow and differentiable programming. It is used for both research and production at Google.
- **HBase:** Apache HBase is an open source distributed non-relational database written in Java and used with elements of the Apache software foundation's Hadoop suite of big data analysis tools. It represents one of several database tools for the input and output of large data sets that are crunched by Hadoop and its various utilities and resources. Some of the popular features of Apache HBase include some kinds of backup and failover support, as well as APIs for popular programming languages. Its compatibility with the greater Hadoop system makes it a candidate for many kinds of big data management problems in enterprise IT.
- **Cassandra:** Cassandra is a distributed NoSQL database designed to manage vast amounts of structured data across an array of commodity servers. Cassandra boasts a unique architecture that delivers high distribution, linear scale performance, and is capable of handling large amounts of data while providing continuous availability and uptime to thousands of concurrent users. Cassandra delivers very consistent performance in a fault tolerant environment. This makes Cassandra ideal for processing online workloads of a transactional nature, where Cassandra is handling large numbers of interactions and concurrent traffic with each interaction yielding small amounts of data.
- **Zeppelin:** Apache Zeppelin is an open-source, web-based "notebook" that enables interactive data analytics and collaborative documents. The notebook is integrated with distributed, general-purpose data processing systems such as Apache Spark (large-scale data processing), Apache Flink (stream processing framework), and many others. Apache Zeppelin allows you to make beautiful, data-driven, interactive documents with SQL, Scala, R, or Python right in your browser.
- **Acumos:** Acumos AI is a platform and open-source framework that makes it easy to build, share, and deploy AI apps. Acumos standardizes the infrastructure stack and components required to run an out-of-the-box general AI environment. This frees data scientists and model trainers to focus on their core competencies and accelerates innovation. Acumos is

part of the LF AI Foundation, an umbrella organization within The Linux Foundation that supports and sustains open-source innovation in artificial intelligence, machine learning, and deep learning while striving to make these critical new technologies available to developers and data scientists everywhere.

- **DIHIWARE:** is an integrated solution aiming to support both the "Access to" and the "Collaborate with" services, providing access to the latest knowledge, expertise and technology during SMEs Digital Transformation pathways toward piloting, testing and experimenting new digital technologies. The Platform leverages on knowledge-driven services in order to maximizing awareness, hands-on experience, knowledge and maturity levels of the ecosystem. Moreover, it provides also Innovation and Collaboration capabilities, so that meeting demand and offer in a marketplace could be transformed into partnership and co-innovation.
- **IoT Catalogue:** Is a web-based tool that enables to pick & choose IoT solutions; it is a wide repository of knowledge, use cases, contacts, etc. of the Internet of Things. A key purpose of the 'IoT Catalogue' is enabling users to explore IoT solutions based on domain-related Value Propositions and/or ICT Problems described on use-cases defined along applications domains. The 'IoT Catalogue' also enables users to inspect solutions and technologies from other domains that might fit to their intents or analyze use-cases similar to their projects thus promoting synergies and reusability between application domains.
- **TEKTON:** Tekton is a powerful and flexible open-source framework for creating CI/CD systems, allowing developers to build, test, and deploy across cloud providers and on-premise systems.
- **ETCD:** is a strongly consistent, distributed key-value store that provides a reliable way to store data that needs to be accessed by a distributed system or cluster of machines. It gracefully handles leader elections during network partitions and can tolerate machine failure, even in the leader node.
- **Superset:** Superset is a modern, enterprise-ready business intelligence web application, fast, lightweight, intuitive, and loaded with options that make it easy for users of all skill sets to explore and visualize their data, from simple line charts to highly detailed geospatial charts.
- **WSO2:** is an open-source technology provider founded in 2005. It offers an enterprise platform for integrating application programming interfaces (APIs), applications, and web services locally and across the Internet. WSO2 offers a platform of middleware products for agile integration, application programming interface (API) management, identity and access management, and smart analytics.
- **Ranger:** is a framework to enable, monitor and manage comprehensive data security across the Hadoop platform. The vision with Ranger is to provide comprehensive security across the Apache Hadoop ecosystem. With the advent of Apache YARN, the Hadoop platform can now support a true data lake architecture. Enterprises can potentially run multiple workloads, in a multi-tenant environment. Data security within Hadoop needs to evolve to support multiple use cases for data access, while also providing a framework for central administration of security policies and monitoring of user access.
- **Accumulo:** is a highly scalable sorted, distributed key-value store based on Google's Bigtable. It is a system built on top of Apache Hadoop, Apache ZooKeeper, and Apache Thrift. Written in Java, Accumulo has cell-level access labels and server-side programming mechanisms. According to DB-Engines ranking, Accumulo is the third most popular NoSQL wide column store behind Apache Cassandra and HBase.
- **KNOX**: is an Application Gateway for interacting with the REST APIs and UIs of Apache Hadoop deployments. The Knox Gateway provides a single access point for all REST and HTTP interactions with Apache Hadoop clusters.

- **Snyk:** is an open-source security platform designed to help software-driven businesses enhance developer security. Snyk's dependency scanner makes it the only solution that seamlessly and proactively finds, prioritizes and fixes vulnerabilities and license violations in open-source dependencies and container images.

4 Conclusions and Next Steps

This deliverable has provided two main results: the AI REGIO Reference Architecture and the Reference Implementation.

The AI REGIO Reference Architecture has been defined starting from the analysis of relevant Reference Architectures for Industrial Internet of Things and I4MS past initiatives. The design phase has leveraged on several aspects:

- The Artificial Intelligence integration, starting from the low layer where the data is processed and analysed, to the higher level where AI supports the services.
- The Data Sovereignty Implementation, supporting data spaces where factories can access for exchange data in a trusted way with the scope of increasing the production quality and the supply chain improvements by the tools, platforms and services offered.
- Requirements collected in the WP2 context related to experiment clusters, in terms of needs, applications and constraints.
- BDVA/DAIRO Grand Challenges scenarios, in order to cover a wide range of use cases (i.e., Smart Product, Smart Factory, Smart Supply Chain).

The AI REGIO Reference Implementation has been proposed, starting from the MIDIH Reference implementation, taking into account the integrations due to the tools for AI supporting the services. At the same time the four clusters (Factory Efficient and Sustainable Manufacturing, Quality Control and Predictive Maintenance, Product Engineering and Lifecycle Management, Robotics and Human Interaction) will feed the platform. The proposed solution is composed of open-source elements avoiding vendor lock-in, lowering costs and barriers for developers, interoperability, and user adoption.

5 References

[1] Deloitte, "AI Enablement on the Way to Smart Manufacturing", Deloitte survey on AI Adoption in Manufacturing, 2020, https://www2.deloitte.com/cn/en/pages/consumer-industrial-products/articles/ai-manufacturing-application-survey.html

[2] Capgemini Research Institute, "Data Powered Enterprise, Why organizations must strengthen their data mastery", 2020, https://www.capgemini.com/wp-content/uploads/2020/11/Data-powered-enterprise-report.pdf

[3] International Data Space Association, "Design Principles for Data Spaces", Position Paper, Version 1.0, p. 39, April 2021, https://design-principles-for-data-spaces.org/

[4] Big Data Value Association, "European Big Data Value Strategic Research and Innovation Agenda", Position Paper, Version 4.0, October 2017, https://www.bdva.eu/sites/default/files/BDVA_SRIA_v4_Ed1.1.pdf

[5] REFERENCE ARCHITECTURE MODEL Version 3.0 April 2019, available at https://internationaldataspaces.org/wp-content/uploads/IDS-Reference-Architecture-Model-3.0-2019.pdf

[6] Dr. Karsten Schweichhart Leader, "Reference Architectural Model Industrie 4.0 (RAMI 4.0)", available at https://ec.europa.eu/futurium/en/system/files/ged/a2-schweichhart-reference_architectural_model_industrie_4.0_rami_4.0.pdf

[7] The Industrial Internet of Things Volume G1: Reference Architecture Version 1.9 June 19, 2019, available at https://www.iiconsortium.org/pdf/IIRA-v1.9.pdf

[8] Functional and Modular Architecture of MIDIH CPS/IOT System v1, available at https://midih.eu/documents/MIDIH%20Reference%20architecture.pdf

[9] Nath S., Stackowiak R., Romano C., 2017. "Architecting the Industrial Internet", Packt Publishing.

[10] Architecture Alignment and Interoperability: An Industrial Internet Consortium and Plattform Industrie 4.0 Joint Whitepaper, IIC:WHT:IN3:V1.0:PB:20171205. February 2018, available at http://www.iiconsortium.org/pdf/JTG2_Whitepaper_final_20171205.pdf