CEN/WS XXX

Date: 2025-XX-XX

prCWA XXXXX:2025

Secretariat: UNE

Reference Architecture for AI solutions' application within process industry – the EU project s-X-AIPI experience

Referenzarchitektur für die Anwendung von KI-Lösungen in der Prozessindustrie – Erfahrungen aus dem EU-Projekt s-X-AIPI

Architecture de référence pour l'application de solutions d'IA dans l'industrie de transformation - l'expérience du projet européen s-X-AIPI

Contents

Page

Cover page

This CEN and CENELEC Workshop Agreement is an agreement, developed and approved by an open independent workshop structure within the framework of the CEN-CENELEC system.

This CEN and CENELEC Workshop Agreement reflects the agreement of the registered participants responsible for its content, who decided to develop this document in accordance with the specific rules and practices available in CEN-CENELEC for the development and approval of CEN/CENELEC Workshop Agreements.

This CEN and CENELEC Workshop Agreement can in no way be held as being a European Standard (EN) developed by CEN and CENELEC, as it does not represent the wider level of consensus and transparency required for a European Standard (EN). Furthermore, it is not intended to support legislative requirements or to meet market needs where significant health and safety issues are to be addressed. For this reason, CEN and CENELEC cannot be held accountable for the technical content of this CEN and CENELEC Workshop Agreement, including in all cases of claims of compliance or conflict with standards or legislation.

The Workshop parties who drafted and approved this CEN and CENELEC Workshop Agreement, the names of which are indicated in the Foreword of this document, intend to offer market players a flexible and timely tool for achieving a technical agreement where there is no prevailing desire or support for a European Standard (EN) to be developed.

The copyright of this document is owned by CEN and CENELEC, and copy of it is publicly available as a reference document from the national standards bodies of the following countries: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Republic of North Macedonia, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey and the United Kingdom.

European foreword

This CEN and CENELEC Workshop Agreement has been developed in accordance with the CEN-CENELEC Guide 29 “CEN/CENELEC Workshop Agreements – A rapid prototyping to standardization” and with the relevant provisions of CEN/CENELEC Internal Regulations - Part 2. It was approved by a Workshop of interested party representatives on 2025-02-14, the constitution of which was supported by CEN and CENELEC following the public call for participation made on 2025-01-07. This CEN and CENELEC Workshop Agreement does not necessarily include all relevant stakeholders.

The final text of the CEN Workshop Agreement is provided to CEN for publication on 2025-03-xx.

Results incorporated in this CWA received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement No 101058715.

The following organizations and individuals developed and approved this CEN Workshop Agreement:

- Daniel Gómez (Chairman), Aníbal Reñones – CARTIF (Spain)
- Sabrina Verardi, Marta Calderaro, Mattia Giuseppe Marzano, Angelo Marguglio - ENG (Italy)
- Walter Quadrini, Francesco Alessandro Cuzzola - POLIMI (Italy)
- Andrea Fernández - AIMEN (Spain)
- Elizabeth Fink, Michael Pölzl, Josef Tausendschön – RCPE (Austria)
- Manuel Alcalá, Ignacio Alberto Fontal Patac - DEUSER (Spain)
- Zeinab Kargar, Petri Kannisto, Bernd Kleimt – BFI (Germany)
- Nenad Stojanović – NISSA (Serbia)
- Steven Carter, CEN/TC 310
-
- Amanda Suo, UNE, Spanish Association for Standardization (Spain)

Attention is drawn to the possibility that some elements of this document may be subject to patent rights. CEN-CENELEC policy on patent rights is described in CEN-CENELEC Guide 8 “Guidelines for Implementation of the Common IPR Policy on Patent”. CEN shall not be held responsible for identifying any or all such patent rights.

Although the Workshop parties have made every effort to ensure the reliability and accuracy of technical and nontechnical descriptions, the Workshop is not able to guarantee, explicitly or implicitly, the correctness of this document. Anyone who applies to this CEN Workshop Agreement shall be aware that neither the Workshop, nor CEN, can be held liable for damages or losses of any kind whatsoever. The use of this CEN Workshop Agreement does not relieve users of their responsibility for their own actions, and they apply this document at their own risk. The CEN Workshop Agreement should not be construed as legal advice authoritatively endorsed by CEN/CENELEC.

Introduction

Impact of Artificial Intelligence in Industrial Settings: Artificial Intelligence (AI) is increasingly recognized as a pivotal force in the industrial digital revolution, complementing advancements in data handling and robotics. AI applications, characterized by simplified interfaces, are designed to be robust and maintainable without the need for a highly specialized workforce. This accessibility extends the functional life of AI applications and reduces the expertise required for their operation.

Objective of the s-X-AIPI Project: The s-X-AIPI project endeavors to research, develop, test, and validate a bespoke suite of trustworthy self-X AI technologies tailored for process industries. This initiative aims to bridge the gap between AI capabilities and traditional automation processes, ensuring that AI tools are both accessible and effective across various industrial applications. The project's core objectives include (Figure 1 represents the s-X-AIPI main objective concept):

- Providing state-of-the-art AI-based sustainability tools to existing process industries and their workforce.
- Enhancing the longevity and user-friendliness of AI applications to minimize reliance on specialized technical skills.
- Deploying trustworthy AI technologies effectively within process industries.

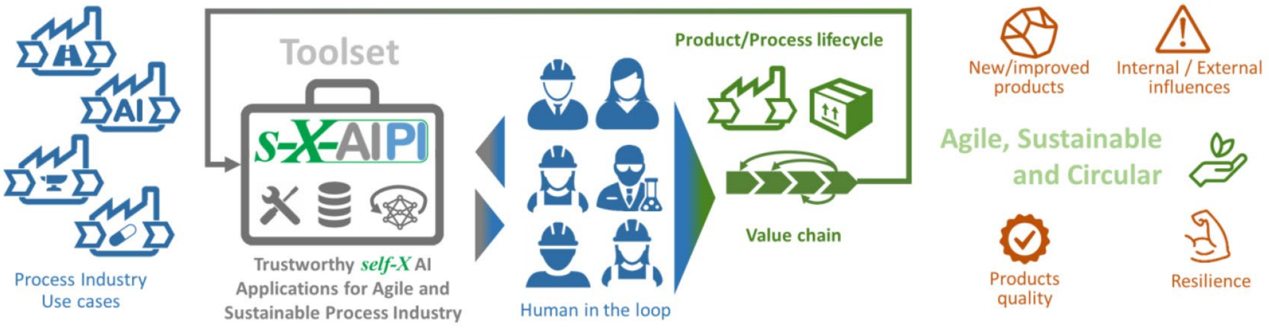


Figure 1 - s-X-AIPI main objective concept

Innovative AI Technologies and Methodologies: At the heart of this project lies the integration of AI with an Autonomic Manager, utilizing the MAPE-K framework (Monitoring, Analyzing, Planning, Execution over a shared Knowledge base) to foster the development of self-improving AI systems. This approach facilitates a practical "learning by doing" model, where continuous adaptations enhance the system's efficacy in real-time applications. The project is distinguished by its incorporation of an AI data pipeline equipped with autonomic computing capabilities, designed to support four realistic use cases in the process industry sectors of asphalt, steel, aluminum, and pharmaceuticals.

Worker-Centric AI Development: A significant aspect of the s-X-AIPI toolset is its focus on accommodating the diverse skill levels of workers, integrating self-adaptation capabilities that respect and enhance the human-in-the-loop role. This approach ensures that the AI technologies developed are not only advanced but also aligned with the practical needs and profiles of the workforce involved.

Expected Outcomes: The primary outcome of the s-X-AIPI project is to cultivate a portfolio of AI technologies that are:

- Trustworthy and integrated into an open-source toolkit for widespread industrial and research application.
- Autonomous, minimizing the need for human intervention in the development and operational processes.
- Broadly integrated across actual process industry value chains, demonstrating the versatile applicability of the developed technologies.

Connection to CWA Objectives: This CWA aims to outline a Reference Architecture for AI solutions tailored to process industry settings, derived from the foundational technologies and methodologies developed under the s-X-AIPI project. It provides a structured approach to designing AI systems that are both innovative and aligned with industry standards, thereby enhancing the capabilities and reach of AI applications in energy-intensive industries.

This CWA has been promoted by the S-A-XIPI project (*'self-X Artificial Intelligence for European Process Industry digital transformation'*). It is a 3-year (May 2022 – April 2025) project funded by the European Union's Horizon Europe Framework Programme for Research and Innovation under Grant Agreement No 101058715. It brings together forteen partners from 6 different European countries (Spain, Italy, Germany, Austria, Serbia and Greece).

1 Scope

This CEN Workshop Agreement (CWA) articulates the preliminary design of the Reference Architecture, specifically tailored for deploying AI-enabled self-X solutions within the Process Industry. This document delineates the architecture's framework and demonstrates its application across diverse industrial scenarios, emphasizing its adaptability and potential for broad implementation.

The primary objectives of this CWA include:

- Provide a comprehensive framework for AI-enabled self-X technologies in process industries.
- Promote the integration of advanced autonomic management systems.
- Facilitate the adoption of AI solutions that are both innovative and compliant with European standards.

A significant focus is placed on contextualizing the MAPE-K methodology within the s-X-AIPI Reference Architecture. This includes a detailed exploration of how this methodology underpins the autonomic features of the AI systems, enabling self-adjustment and improved decision-making capabilities within industrial operations.

An extensive analysis of relevant Reference Architectures is provided, including established frameworks like RAMI 4.0, IIRA, and emerging technologies such as FIWARE and IDS RAM 4.0. The document also examines newly developed architectures in the European context, like BEinCPPS and CAPRI, highlighting their relevance and integration into the s-X-AIPI architecture.

Applicability: While the Reference Architecture is designed to be versatile and generic, sufficient attention is given to its application in selected industrial scenarios. This ensures that the architecture not only supports a wide range of applications but also meets specific industry needs, facilitating tailored adaptations where necessary.

The scope of this CWA does not encompass the definition of safety-related requirements. Additionally, this document is intended to be informative, aimed at augmenting existing standards rather than replacing or simplifying mandatory production procedures. It seeks to provide a structured approach to integrating AI technologies in process industries, enhancing operational efficacy and innovation without compromising established procedural standards.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/TS 15066:2016 apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- ISO Online browsing platform: available at https://www.iso.org/obp
- IEC Electropedia: available at https://www.electropedia.org/

4 Background and General Vision

4.1 Project Objectives

The impact of Artificial Intelligence (AI) is highly recognized as a key driver of the industrial digital revolution together with data and robotics. AI tools applications are needed with more simplified interfaces without requiring highly skilled workforce but exhibiting longer useful life and requiring less specialized maintenance.

The overall objective of s-X-AIPI project is to research, develop, test and experiment an innovative toolset of custom trustworthy self-X AI technologies and applications. More specifically technical objectives are:

- Provide existing process industries and its workers state of the art AI-based sustainability tools.
- Longer useful life of AI apps, deployment more simplified interfaces, without requiring highly skilled workforce, and requiring less specialized maintenance (so applications with minimal human expert intervention).
- Achieving an effective deployment of trustworthy AI technologies within process industries, close the gap between AI and automation.

Within this project Self-X is the combination of the AI as the intelligent processing system and an Autonomic manager based on MAPE-K (continuous Monitoring-Analyzing-Planning-Execution flow based on the Knowledge of the AI system under control) for developing self-improving AI systems. This will be realized by an adaptation loop, which enables “learning by doing” using MAPE-K model and self-X abilities as proposed by autonomic computing.

In this way, s-X-AIPI toolset of AI technologies are intended to include an innovative AI data pipeline with autonomic computing capabilities (self-X AI and Autonomic Manager), architecture, realistic datasets together with their respective algorithms derived from the demonstration in four realistic use cases of process industry (asphalt, steel, aluminum and pharmaceutics). Moreover, s-X-AIPI technologies will consider workers’ heterogeneous skill levels and self-adaptation capabilities to the actual profile of the worker respecting their human-in-the-loop role. Figure 1 represents the idea of the project concept.

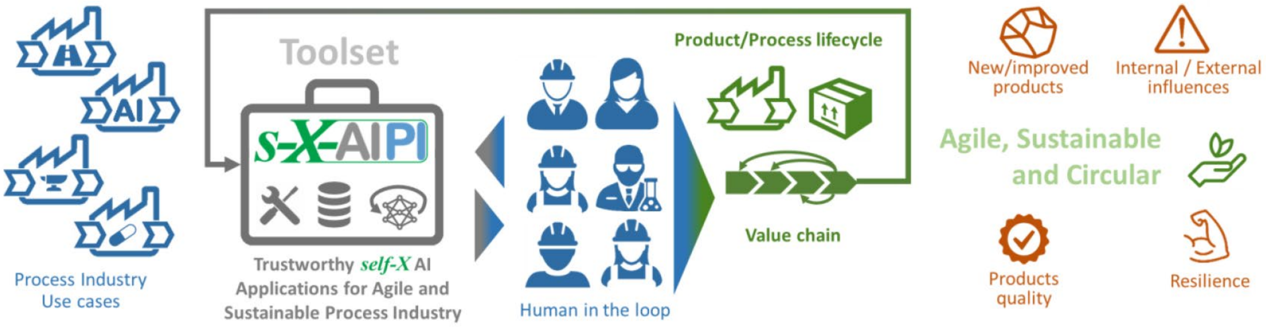


Figure 2 - s-X-AIPI main objective concept

The main expected result is to generate a showcase portfolio of AI technologies (data sets, AI model and applications):

- Trustworthy integrated into an innovative Open Source toolset available for industry and research.
- Autonomous with self-adaptation capabilities to the actual profile of the worker, minimizing human involvement in the AI application development loop.
- Integrated in 4 actual process industries value chains.

Additionally, this project has a result [RES6] directly related with AI for Process Industry Reference Architecture, which express that s-X-AIPI architecture build upon stablished FIWARE foundation to create an open implementation to support the orchestration of complex data flows defined in the self-X AI data pipeline, where the AI prototype applications will be developed and tested.

4.2 The MAPE – K Methodology

The **MAPE-K (Monitor-Analyze-Plan-Execute over a shared Knowledge)** methodology implements a feedback loop as reference control model for autonomic and self-adaptive systems. The primary goal of any feedback control system is to remove or reduce humans from the loop, and therefore MAPE-K focuses upon autonomous decision-making and self-adaptation.

The MAPE-K feedback control loop [1] performs Monitor-Analyze-Plan-Execute phases over a shared Knowledge. The phases can be described as following:

- The **Monitoring Phase (M)** is primarily concerned with collecting data from the self-adaptive system and the environment in which it operates. Hardware sensors provide raw data collected, persisted, and used in runtime models to guide subsequent analysis and self-adaptation decisions.
- The **Analysis Phase (A)** determines whether adaptation actions are required, based on the current and predicted state of the system, the environment is operating in, and its defined goals, safety constraints, and quality of service specifications. Automated analysis enables timely and fast reactions to changes in the environment and emergent situations.
- In the **Planning Phase (P)** the machine plans self-adaptation actions such as switching states to perform different tasks, reconfiguring existing features, activating or deactivating sensors, or modifying polling frequencies to preserve power or to collect additional information about the system or its environment.
- During the **Execution Phase (E)**, the previously generated plan or adaptation strategy is executed on the physical machine or device.
- The **Knowledge Base (K)** is a collection of data coming from the managed system and environment, adaptation goals, and other relevant states that are shared by the MAPE components.

Figure 2 shows in states diagram the interaction among the different phases described above. The Perceptors observe the real world and transfer the information to the Monitor phase detecting the anomalies. At the same time the simulator component compute from the current information which are the expectations, in order to allow the Planner to take a decision with the support of the Diagnoser, referring to the models used. Once the decision is taken, the Executor triggers the action must be done by the effectors in the real world.

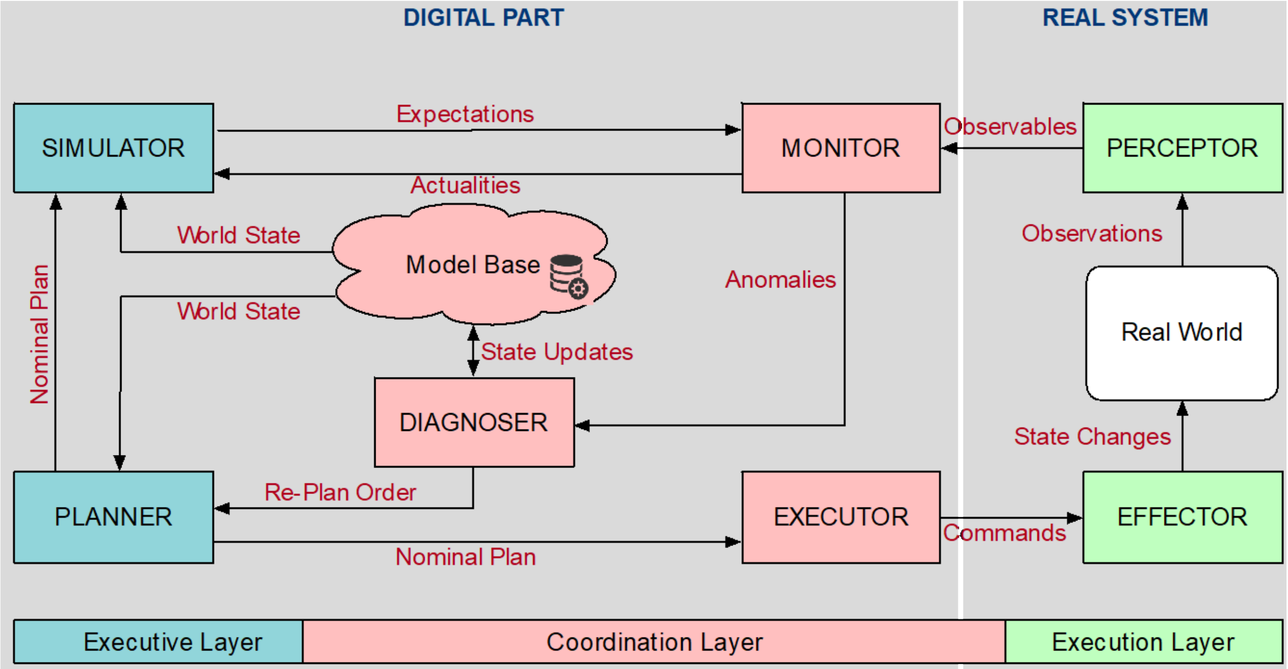


Figure 3 - MAPE-K Methodology state diagram

The analysis done about the MAPE-K involved also the cognitive computing concept, an approach that is based on recent research work on Artificial Intelligence providing methods to construct and operate systems that “know what they are doing” (like human experts), to support the proper application of MAPE-K in s-X-AIPI Reference Architecture design. This is supported by the exploration of the unknown situations related to the four domains to resolve “new, unforeseen issues” and adapt over time. Since there are always unknown situations that are not clear at the time of design, the monitoring component must respond to this unknown event at the specified time and bring the system to a safe state.

In the same way, it was carried out an investigation about Autonomic Computing (AC), a distributed computing resources with self-managing characteristics, adapting to unpredictable changes, hiding intrinsic complexity to operators and users. The AC system concept is designed to make adaptive decisions, using high-level policies, constantly checking and optimizing its status and automatically adapt itself to changing conditions.

The results obtained from the investigation produced the application of MAPE-K Methodology in s-X-AIPI project, shown in Figure 3, that will support the design of the Reference Architecture defined in section 4.1.

In this case, the real world will interact with the *AI Pipelines*, capable to provide insights and advanced rules on the way to act or react to unknown events. The *Perceptors* collect the data from the plant, external applications and AI pipelines outputs, providing inputs to the *Monitor phase* in the *Autonomic Manager* (AM). Therefore, the *Real System,* identified in Figure 3, is monitored by the *AI Pipelines*, becoming an AI system which makes the desired AI applications able to process the data collected from the Preceptors. The AI Data Pipelines, in conjunction with the data monitored in the real system represents the *Knowledge base* for the Autonomic Manager. It might be influenced by the human-in-the-loop, the Artificial Data and the Labelling applications. The core part of the AM is the *Analyze phase,* in charge of processing the outcomes of the monitoring component with the aim to provide all the information needed to take decisions in the *Plan phase*. At the end of the loop, the *Executor phase* will return the result to act directly to the *AI Pipelines* (and subsequently to the *Real System* through the *Effectors*) or asking for Human intervention.

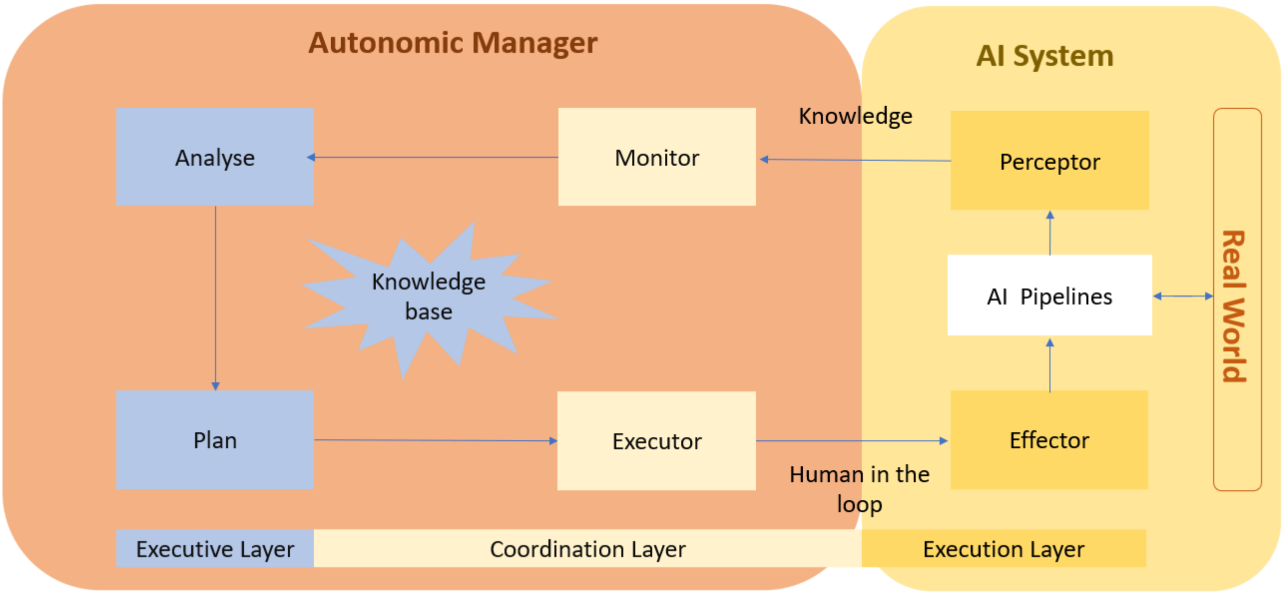


Figure 4 - MAPE-K methodology applied in s-X-AIPI

4.3 AI Data Pipelines

s-X-AIPI project looks to apply a new approach to build AI based applications for the process industry. The approach looks for more autonomous AI applications from the AI creating process itself.

Our project methodology intends to apply an AI data pipeline suitable for AI online operation and autonomic behavior, as depicted in the following Figure 4.

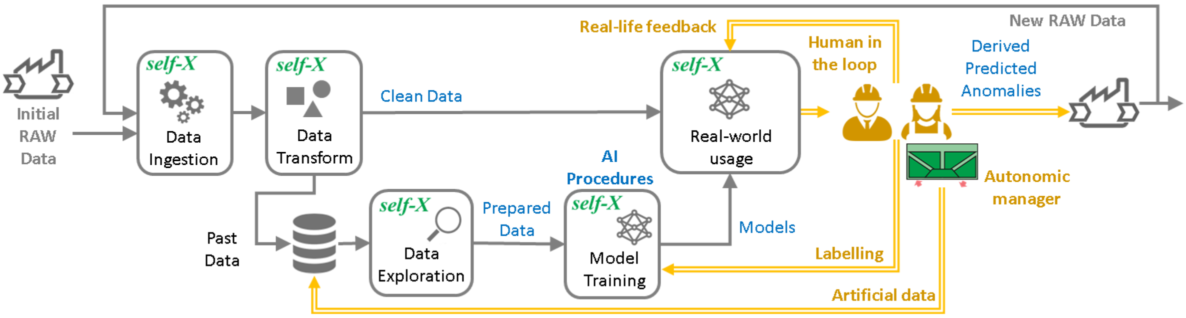


Figure 5 - AI data pipeline

As described in the document [D2.1] of s-X-AIPI, different pipeline components exhibit certain (self-X) abilities like self-Healing or self-Configuration. It means that each of these components will go through MAPE-K process (realized as Autonomic manager, as presented in Figure 3). For some components (that correspond to core AI tasks) there is also ”Human in the loop”, supporting some more challenging activities, like labeling of new datasets (notice that such a request should be outcome from the Autonomic manager).

Therefore, we can state the following:

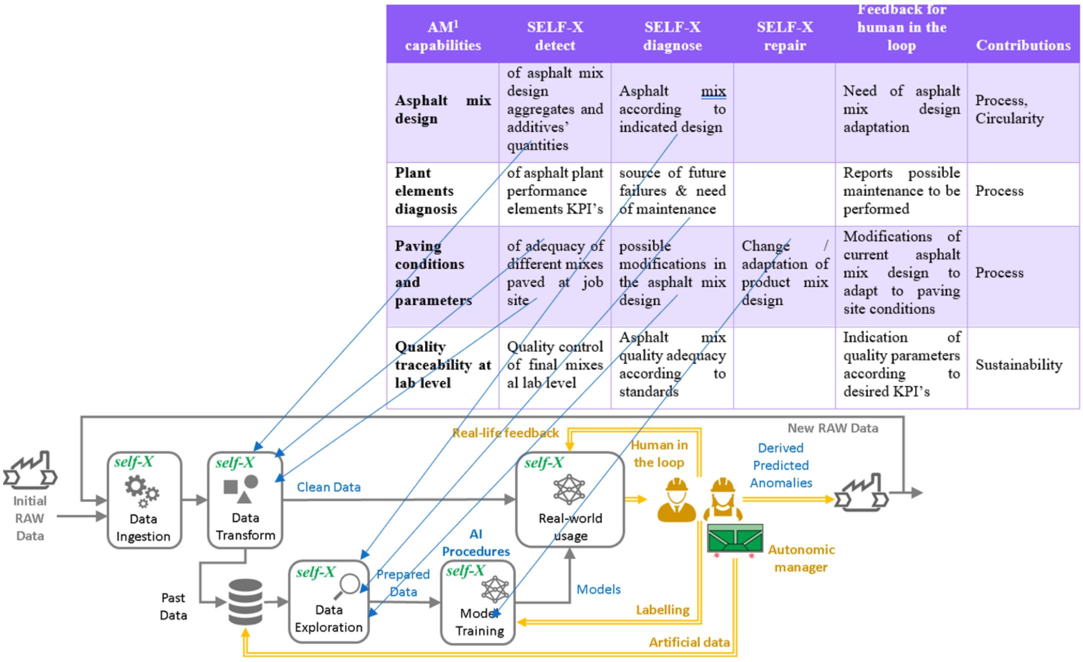
1) AI data pipeline executes AI processing related to the pilots, realizing desired AI applications (e.g. predictive maintenance, anomaly detection) based on the real world (data).

2) each component in the pipeline provides data about the own execution (like in a log system).

3) Perceptor is responsible for offering this data to the Autonomic Manager (e.g. using a Message Bus system).

4) Autonomic Manager is doing corresponding self-x (based on MAPE-K) process, specific for each pilot application/scenario (as illustrated in the Figure 5).

5) The results are sent to Message bus / or executed directly on the AI pipeline, or involving Human in the loop.



AM[1] capabilities	SELF-X detect	SELF-X diagnose	SELF-X repair	Feedback for human in the loop	Contributions
Asphalt mix design	of asphalt mix design aggregates and additives' quantities	Asphalt mix according to indicated design		Need of asphalt mix design adaptation	Process, Circularity
Plant elements diagnosis	of asphalt plant performance elements KPI's	source of future failures & need of maintenance		Reports possible maintenance to be performed	Process
Paving conditions and parameters	of adequacy of different mixes paved at job site	possible modifications in the asphalt mix design	Change / adaptation of product mix design	Modifications of current asphalt mix design to adapt to paving site conditions	Process
Quality traceability at lab level	Quality control of final mixes al lab level	Asphalt mix quality adequacy according to standards		Indication of quality parameters according to desired KPI's	Sustainability

Figure 6 - Mapping of the requirements from a use case on the AI pipeline

5 Relevant Reference Architectures

5.1 FIWARE for INDUSTRY

FIWARE Foundation drives the definition of key open standards that enable the development of portable and interoperable smart solutions in a faster, easier and affordable way.

From a technical perspective, FIWARE brings a curated framework of Open-Source software components which can be assembled together and combined with other third-party platform components to build platforms and smart solutions.

It eases the digital transformation by fostering smart organizations' innovation in multiple application domains - as manufacturing, utilities, etc. – thankful to three main principles:

- Facilitating interoperability and data exchange through a smart usage of data, using data models de facto standardized in combination with Standard APIs for data management and exchange.
- Enabling interoperable software implementations for Context Information Management and Big Data services, avoiding industrial data silos, supporting the connection to IoT and facilitating the implementation at the edge, thankful to standard architectures and Open-Source software components.
- Enabling easy plug&play integration with 3rd party solutions and services, avoiding any vendor lock-in scenario.

FIWARE supports the adoption of its agile approach and the related Open-Source components thankful to a FIWARE marketplace and a related Catalogue, where the portable and interoperable solutions are available to the community. In 2015, the manufacturing domain has been part of the launch of the FIWARE for Industry (F4I) initiative, to help the needs of Manufacturing Industry business scenarios, afterwards evolved in a domain agnostic "Powered by FIWARE" concept.

Any "Powered by FIWARE" platform or solution is based on the presence of a main and mandatory component, the FIWARE Orion Context Broker Generic Enabler, which is capable to manage context information, perform any update and bring access to the context.

As depicted in Figure 6, the FIWARE Context Broker represents a cornerstone with which it is possible to:
- capture updates by the Internet of Things, Robots or 3rd Party systems and translating required actuations, if necessary;

- support usage control and even publish or monetize managed context data;

- support the end-users in their digital transformation path, thankful to the smart behavior of FIWARE applications and their capabilities of advanced processing, analysis and visualization of context information.

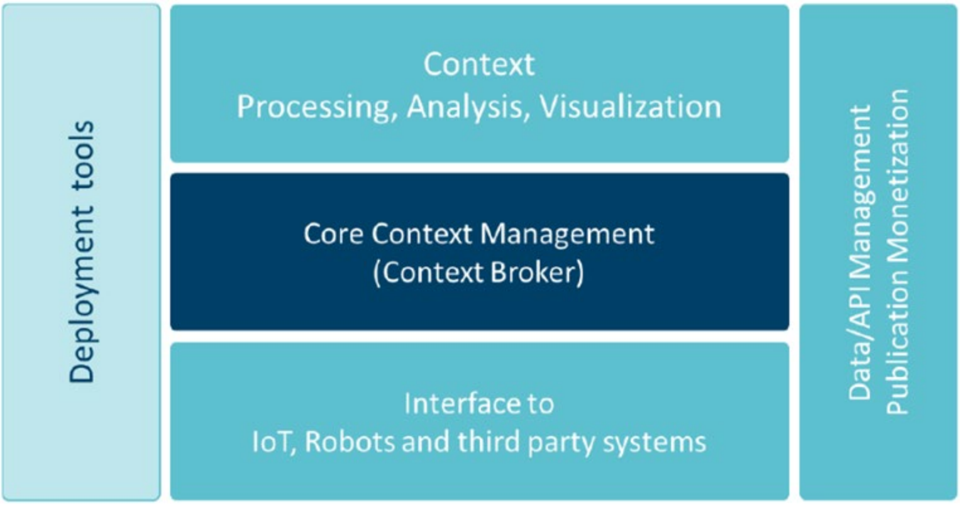
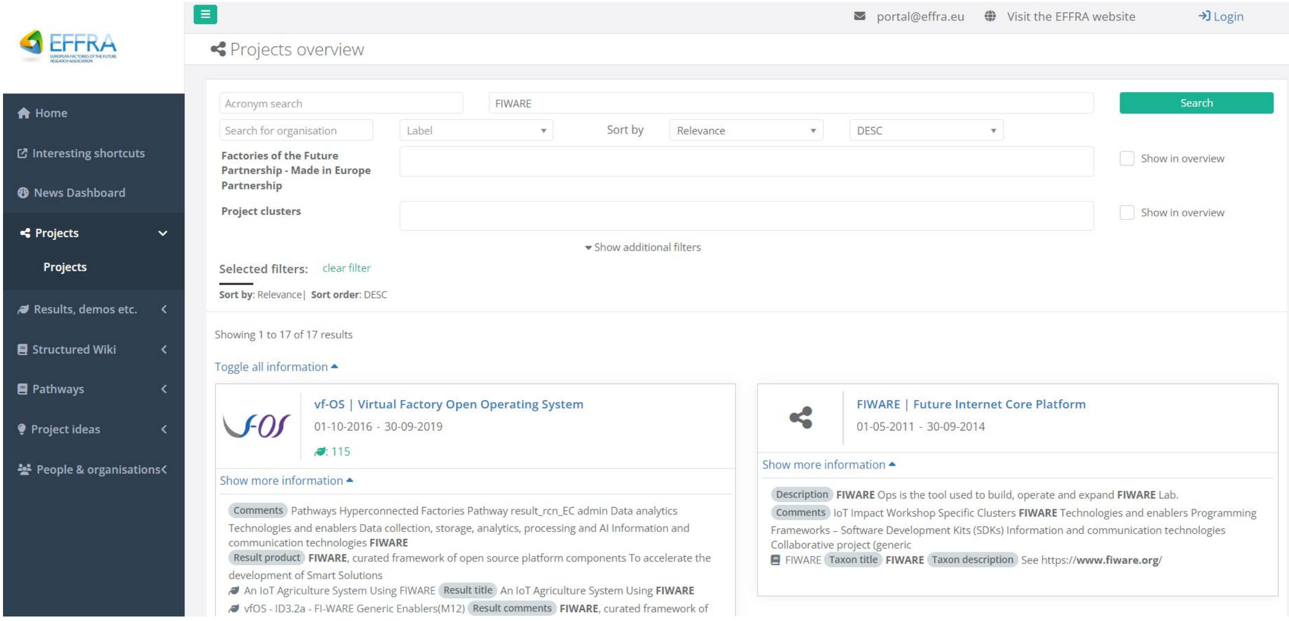


Figure 7 - FIWARE Overall Architecture

Therefore, thankful to FIWARE vision, industrial silos have been limited or either removed, enhancing the development and implementations of Smart-Digital-Virtual Factory scenarios. Those are enabled by the integration of 15 FIWARE Generic Enablers and of additional 16 Generic Enablers currently under incubation, available in the FIWARE Catalogue [2].

The development and successful deployment of the FIWARE solutions have been pursued via a different set of projects and European initiatives. In the context of the FIWARE Marketplace [3], at least 19 applications in the manufacturing industry domain have been reported. In the context of the European Factories of the Future Research Association (EFFRA), around 17 European funded projects have been classified in the Innovation Portal to use or develop "powered by FIWARE" solutions (see Figure 7).



FIWARE

Sort by: Relevance

DESC

Factories of the Future Partnership - Made in Europe Partnership

Project clusters

Selected filters:

Sort by: Relevance| **Sort order**: DESC

Showing 1 to 17 of 17 results

vf-OS | Virtual Factory Open Operating System
01-10-2016 - 30-09-2019
: 115

Comments Pathways Hyperconnected Factories Pathway result_rcn_EC admin Data analytics Technologies and enablers Data collection, storage, analytics, processing and AI Information and communication technologies **FIWARE**
Result product **FIWARE**, curated framework of open source platform components To accelerate the development of Smart Solutions
An IoT Agriculture System Using FIWARE Result title An IoT Agriculture System Using **FIWARE**
vfOS - ID3.2a - FI-WARE Generic Enablers(M12) Result comments **FIWARE**, curated framework of

FIWARE | Future Internet Core Platform
01-05-2011 - 30-09-2014

Description **FIWARE** Ops is the tool used to build, operate and expand **FIWARE** Lab.
Comments IoT Impact Workshop Specific Clusters **FIWARE** Technologies and enablers Programming Frameworks – Software Development Kits (SDKs) Information and communication technologies Collaborative project (generic
FIWARE Taxon title **FIWARE** Taxon description See https://**www.fiware.org**/

Figure 8 - Initiatives implementing FIWARE solutions in EFFRA Innovation Portal

The above picture shows a snapshot of the continuously evolving environment where the Smart Industry Mission Support Committee, chaired by ENGINEERING and NEC, acted to launch and promote the FIWARE Smart Industry Reference Architecture, enabling digital transformation of industrial environment thankful to:

- The integration of information gathered from robots, other machines and sensors in the shop floor, as well as information systems, breaking information silos. This is enabled by Orion Context Broker.
- Media streams from cameras are transformed into valuable context information by Kurento.
- IDAS IoT Agents act as a bridge among sensors and the Context Broker, handling multiple IoT protocols (MQTT, OPC UA, CoAP/OMA-LWM2M, OneM2M, etc.).
- ROS-2 robots are interfaced using Fast RTPS, adopted as default communication middleware in ROS-2.
- Tailor-made system adapters cope with rest of shop floor machines and information systems.
- The NGSI API provides a simple yet powerful RESTful API for getting access to context / Digital Twin data and representation.
- History data is processed using different processing engines (e.g., Hadoop or Flink) to extract insights or derive smart actions. Complex Event Processing, Advanced AI or machine learning functions can be implemented on top of Apache Flink or Spark.
- Business Intelligence is implemented via Apache Superset.
- Operational dashboards are based on the Wirecloud web mashup framework, enabling visualization and dashboarding via Grafana.
- Context Data/API management, publication and monetization supports the trusted data exchange in a Data Space environment.

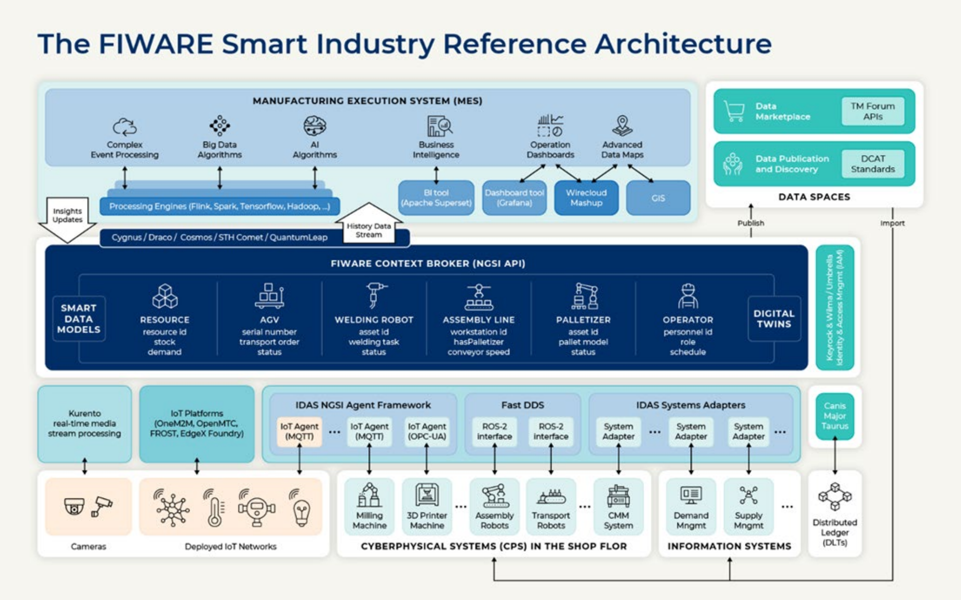


Figure 9 - FIWARE Smart Industry Reference Architecture

History Data streams and insights updates from Manufacturing Execution Systems are at the basis of the data processing, which enables Data in Motion (Industrial IoT) and the Data at Rest (Industrial Analytics) processing. A distributed execution framework to support dynamic processing flows over cloud and edges is also enabled in the FIWARE Smart Industry Reference Architecture, depicted in Figure 8. The s-X-AIPI Reference Architecture will indeed support the two dynamic processing flows enabled by a "powered-by-FIWARE" architecture.

5.1.1 NGSI-LD and Smart Data Model Initiative

The FIWARE Foundation overarching vision based on Open Source principles and interoperable solutions fully supports standardization measures, capable to overcome technological barriers and subsequent industrial data silos.

Among those, the NGSI API fully supports FIWARE vision. The NGSIv2 API, which was defined by members of the FIWARE Community and is currently used in many systems in production within multiple sectors, provides a simple yet powerful RESTful API for getting access to context / Digital Twin data.

Evolution of the API has taken place within the ETSI ISG CIM (Context Information Management Industry Specification Group), where members of the FIWARE Community and the FIWARE Foundation have led the definition of an evolved version of the API, known as the NGSI-LD API, whose specifications were first published by ETSI in 2019 and continue to evolve.

The NGSI-LD API is used as the data integration API and is implemented by the core component of any "powered by FIWARE" architecture via the Context Broker component. Different alternative Open Source implementations of a Context Broker are available within the FIWARE Community, namely the Orion-LD, Scorpio and Stellio products. Furthermore, the NGSI.JS library provides a series of JavaScript functions allowing developers to connect and push context data to any NGSI compliant context broker.

Recently, ETSI ISG CIM provides additions and corrections to the GS-009 NGSI-LD API specification[1] based on feedback received from developers in the linked data, IoT, mobile apps and smart applications communities,

---

[1] An early draft is dated February 2023, 16th, and available at ETSI Work Item "RGS/CIM-009v171".

as well as from end users and stakeholders, supporting the Web of Trust concepts and furtherly emphasizing the Data Space initiatives continuum. Those additions and corrections have been considered in order to support a smooth implementation of the s-X-AIPI Architecture, ensuring potential scalability and future improvement in a long term period [4].

With the advancement on European policy framework on Artificial Intelligence and the subsequent proposal of the regulation called “Artificial Intelligence Act” [5], the European Commission drive a first draft of a standardization request (SR) to the European Standardization Organizations, namely ETSI, CEN and CENELEC, in support of safe and trustworthy artificial intelligence.

In June 2022 ETSI launched the Operational Co-ordination Group on Artificial Intelligence (OCG AI), who will coordinate the comments on the standardization request while the resulting technical work will be performed in the appropriate technical committees. Despite the AI Act is not yet adopted by the European Parliament and Council, the ETSI OCG AI related work should be followed in order to support the development of harmonized standards, and as well to ensure the adoption of standardized measure and technologies in the Reference Architecture implementation.

The Smart Data Models initiative is a second proof of FIWARE approach to standardization [6]. It provides a library of Data Models described in JSON/JSON-LD format which are compatible respectively with the NGSIv2/NGSI-LD APIs or would be useful for defining other RESTful interfaces for accessing Digital Twin data. This further initiative, joined by several organizations like TM Forum, OASC or IUDX, fully supports interoperability purposes.

5.2 Reference Architectural Model Industrie 4.0 (RAMI 4.0)

The concept of the so-called “Industry4.0” has been introduced for the first time during 2011 Hannover Messe [7].

The idea behind it is to enable significantly more interaction between shop-floor devices and high-level business systems, building on the previous generation of industrial monitoring and control systems. As a result, production flexibility is increased, and the International Society of Automation 95 (ISA 95) [8] traditional rigorous hierarchical approach has changed [9]. This hierarchy is shown in the pyramid of Figure 9 and the structure expresses the typical model of production systems automation before the Industry4.0 paradigm, when each level of the pyramid relies on the layer below, with the lower levels providing the data and control necessary for the higher levels to monitor, make decisions, and take action. This hierarchical structure allows a clear separation of responsibilities and enables the integration of different systems and technologies [10]. The pyramidal layers are built on a "layer 0", which includes the hardware (sensors, actuators) involved in the production process that are responsible for the direct control and the collection of data from the field devices, and then build up layer 1 the Programmable Logic Controller (PLC) layer which is responsible for the control of the equipment, the collection of data, and the monitoring of alarms and events, layer 2 (SCADA layer), layer 3 (Manufacturing Execution System, MES, layer) and layer 4 (Enterprise Resource Planning, ERP, layer). The IEC 62264 is an international standard established by the International Electrotechnical Commission, which serves as a complementary standard to ISA 95. It provides a comprehensive framework for the integration and communication of enterprise and control systems in the context of Industry 4.0. IEC 62264 provides a standardized approach to integration and communication between different systems within a manufacturing organization since it defines a set of functional requirements and protocols for communication between these systems, which is critical for effective industrial control and automation [10].

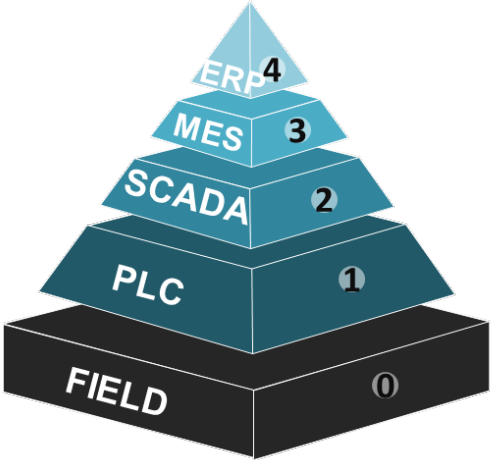


Figure 10 - The “automation pyramid” [11]

ISA-95 automation system alone is not able to address the dynamics of a flexible production environment, and this can be seen by the fact that many businesses are already requesting new models and architectures for dynamic and digitized production as they struggle with high costs when making even simple adjustments to their ISA-95-based production automation systems [12]. In particular, there is a transition from legacy automation technology as defined by ISA-95 to highly distributed Internet of Things (IoT)- based automation systems that fully utilize Internet technologies, thus enabling the implementation of the so-called “Reference Architectural Model Industrie 4.0” (RAMI 4.0) [12], a reference model that provides a common understanding and a common language for Industry 4.0, defining the key components, their relationships and interactions required for Industry 4.0 and outlining the technical requirements and standards needed to implement Industry 4.0 solutions.

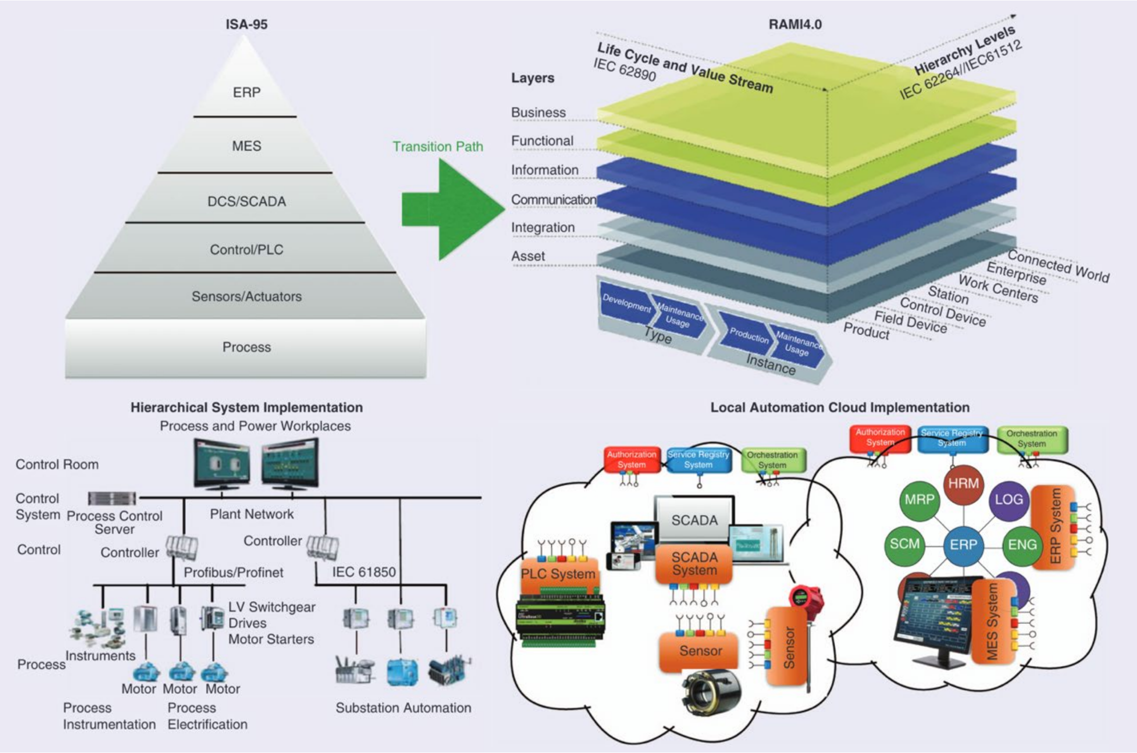


Figure 11 - Digitalization trend towards RAMI 4.0

This new paradigm generates value through repurposing and extracting useful information from raw manufacturing data enabling the possibility of data's flow between different elements of a more complicated system that may connect directly with one another and with their local environment on many levels. This new representation (Figure 10) emphasizes, from a network perspective, how close each layer is to the field, enabling data collection straight from it. In order to monitor and control the production system and eventually support a decision-making process through the computation of Key Performance Indicators, the layers 1 to 4 require these data (KPIs) [11].

RAMI 4.0 utilizes a three-dimensional coordinate system to encompass all crucial elements of Industry 4.0. The three axes allow for the mapping of all critical aspects of Industry 4.0, creating a 3D representation of Industry 4.0 solutions. It serves as a guide for aligning the needs of various fields with various standards, to implement Industry 4.0 methods, and enables the recognition of overlapping requirements, gaps and their resolution in accordance with related industrial standards [13].

Layers in the vertical axis (Figure 11) indicate the appearance from various angles (a look from the market aspect, a look from a perspective of functions, information, communication, a look from an integration ability of the components) [14].

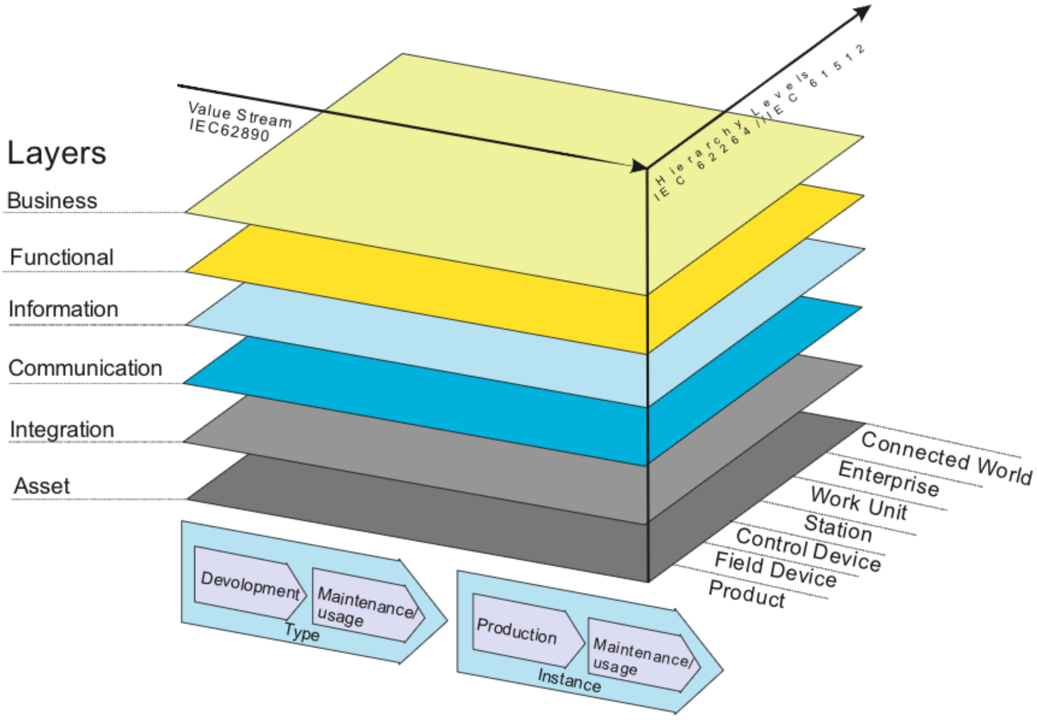


Figure 12 - RAMI 4.0 [13]

Moreover, the product life cycle and the value stream it comprises are crucial criteria in contemporary engineering and this characteristic can be seen along the left-hand horizontal axis. Examples given include continuous data collection throughout the life cycle. Even the complete digitalization of the entire development-market chain offers enormous possibilities for improving goods, machinery, and other Industry 4.0 architecture layers over their entire lifecycle [15].

The following model axis (right on the horizontal level) describes the functional positioning of the Industry 4.0 components. The functionality of the components is specified in this axis; there is no implementation specification, only the function assignment.

In 2016, Zezulka et al [15] explained in detail the layers:

1. The asset layer, which is the bottom layer of RAMI 4.0, is in charge of fostering a shared knowledge of the physical assets present in an Industry 4.0 system. These resources could be anything, from infrastructure and buildings to tools and machinery. The asset layer is used to collect and store

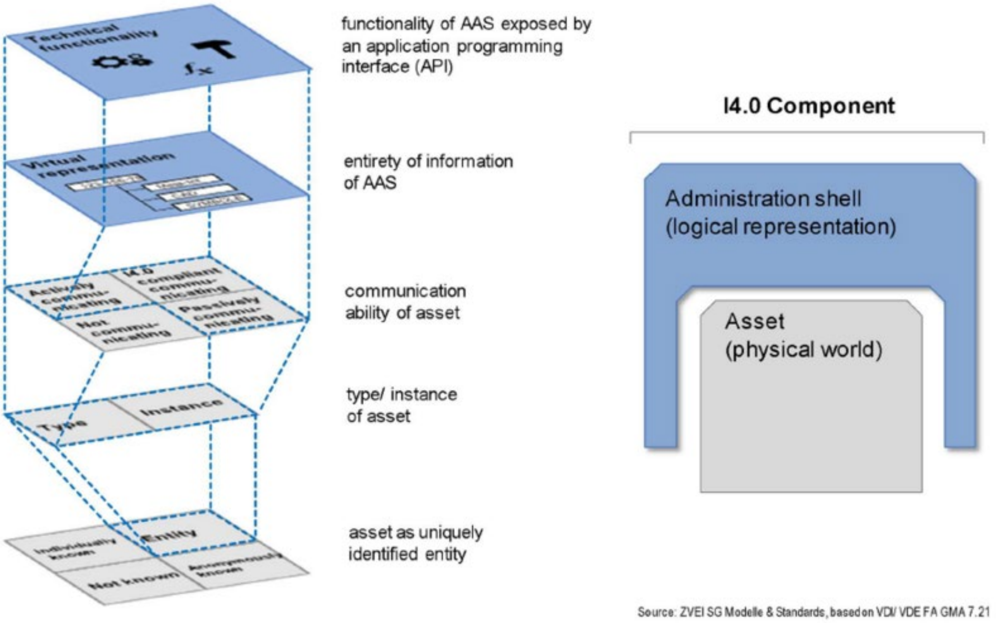
information about these assets' state, performance, and maintenance requirements in addition to providing a standard method of defining and identifying them.

2. Integration layer: In an Industry 4.0 environment, this layer is in charge of linking and coordinating the various systems and devices. It ensures that the various systems and devices can cooperate and allows data and information to flow between the various architectural layers. The integration layer also makes it possible to integrate older Industry 4.0 technologies with legacy systems and equipment.
3. Communication layer: This layer offers the infrastructure and protocols required for inter-system and inter-device communication. It facilitates the safe and dependable transfer of data and information between the various architectural layers and supports wired and wireless networks, among other communication technologies.
4. The information layer is in charge of managing and processing the data and information that the other levels acquire and produce. In addition to enabling the production of a digital twin of the physical assets and processes, it also comprises tools and technology for data storage, analysis, and visualization.
5. Functional layer: This layer offers the capabilities and functionality needed to regulate and improve the Industry 4.0 environment. It facilitates the development of intelligent, self-optimizing systems and encompasses technologies such as automation, robotics, and artificial intelligence.
6. Business layer: This layer is in charge of giving the Industry 4.0 environment its business context and objectives. It covers the operational procedures and procedures that the other levels make possible, such as supply chain management, customer relationship management, and production planning. The managerial and decision-making abilities necessary for the business to function are also included in the business layer.

5.2.1 Asset Administration Shell

The Asset Administration Shell (AAS) is defined as the standardized digital representation of an asset, which might be considered as any physical or logical object that has an actual or perceived value for an organization (IEC 62832-1:2020 [16]). Together, an asset and its corresponding AAS form the building blocks of the I4.0-component, which can represent any component within the RAMI 4.0 hierarchy levels, and it is the core element of Industry 4.0.

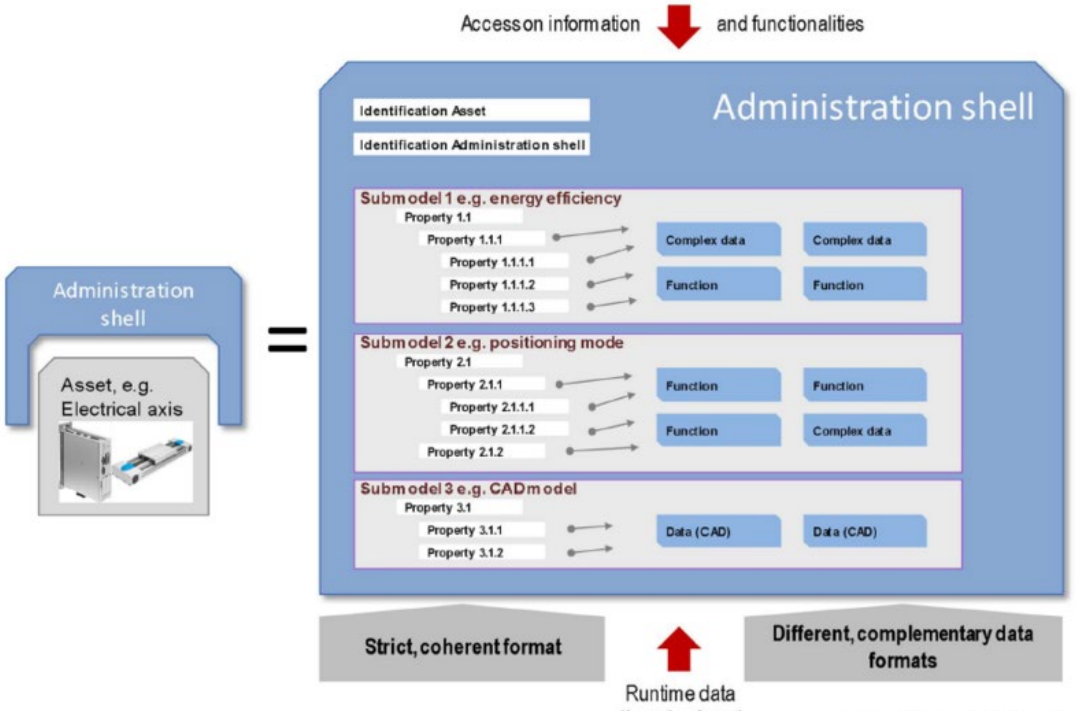
The AAS provides access to the entirety of information of its asset and ensures the interoperability between I4.0 components in modern Industrie 4.0 systems by managing the external communication of the asset with the connected, digital world in a standardized format. Ultimately, the AAS ensures the unique, unambiguous identification of its asset via unique identifiers to be addressable within the network, acting as a standardized and secure communication interface of the asset in the digital world as shown in Figure 12.



Source: ZVEI SG Modelle & Standards, based on VDI/ VDE FA GMA 7.21

Figure 13 - AAS common layers and logical view

The AAS metamodel is a manufacturer-independent standardization that defines structural principles of the AAS in a formal manner (UML) with the aim of enabling information exchange via Industrie 4.0 communication protocols. The structure of the AAS usually consists of a number of domain-specific submodels that encapsulate all the information and functionalities of the asset, including its properties, capabilities, and references and relationships to other components or objects, among others. As shown in Figure 13, the information within the AAS follows a tree-based structure of properties in which submodels support the integration of different data sources of the asset, enabling the integration of external sources with proprietary interfaces by acting as reflexive interfaces or common points of interaction.



Source: ZVEI SG Modelle & Standards

Figure 14 - AAS exemplary information structure

The AAS is supported by multiple serializations mappings that include XML, JSON, or OPC-UA. According to its behavior, three types of AASs can be distinguished: type I, passive AASs represented as serialized files containing static information; type II, reactive AASs existing as runtime instances; and type III, proactive AASs that show active behavior i.e., communication via standardized Industrie 4.0. language. On this line, the information model of the AAS support the integration of concept descriptions and semantic identifiers in the different elements of the AAS e.g., via ECLASS semantic identifiers, enabling the transition from implicit to explicit semantics in the communication between AASs in the connected, digital world.

5.3 Big Data Value Association (BDVA)

The European BDVA (Big Data Value Association) has designed a BDV Reference Model that serves as a common reference framework to locate Big Data technologies on the overall IT stack.

The Strategic Research and Innovation Agenda (SRIA) define the overall goals, main technical and non-technical priorities, and a research and innovation roadmap for the European contractual Public Private Partnership (cPPP) on Big Data Value.

The BDV Reference Model has been developed considering input from technical experts and stakeholders along the whole Big Data Value chain, as well as interactions with other related PPPs, in order to create a common understanding between all members interested in Big Data and Artificial Intelligence applications applied to the Smart Manufacturing Industry domain. It is structured into horizontal and vertical concerns which distinguishes between elements that are at the core of the BDVA and the features that are developed in strong collaboration with related European activities.

The Big Data Value Reference Model is shown in Figure 14.

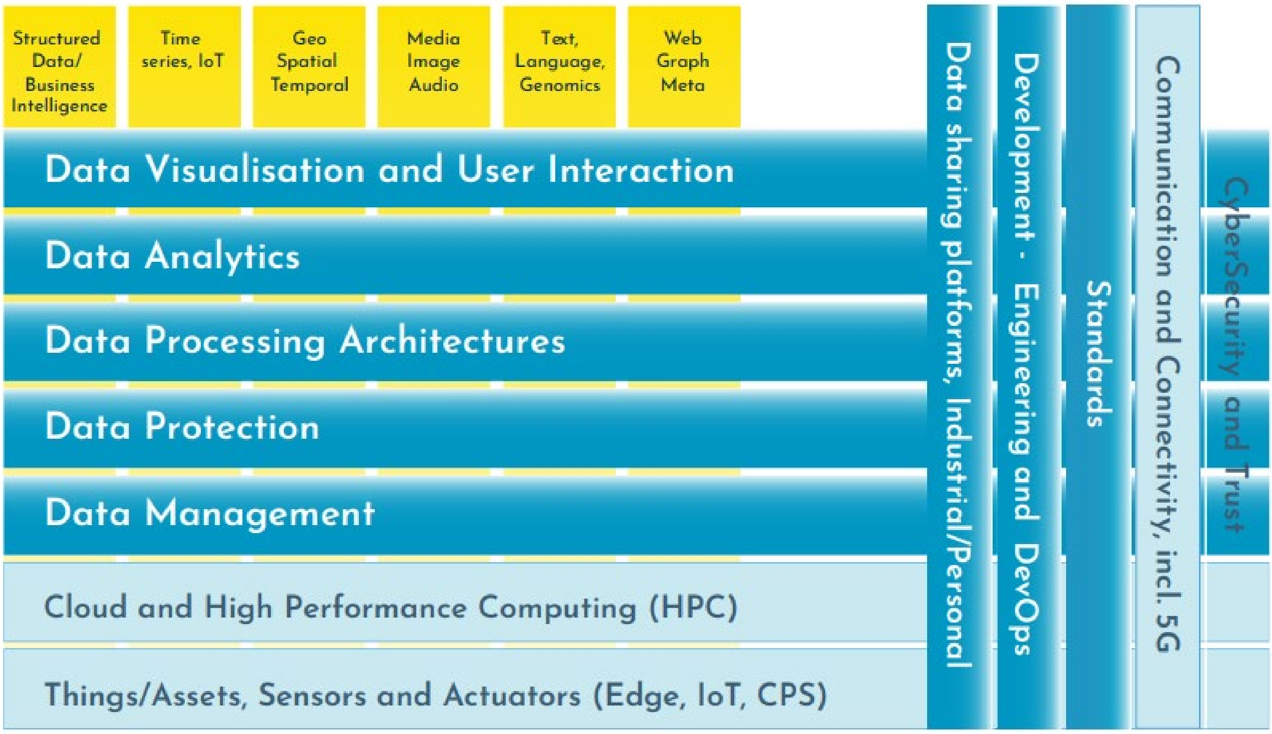


Figure 15 - Big Data Value Reference Model

The horizontal concerns cover specific aspects along the data processing chain, starting with data collection and ingestion, and extending to data visualization. It should be noted that the horizontal concerns do not imply a layered architecture. As an example, data visualization may be applied directly to collected data (the data management aspect) without the need for data processing and analytics.

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

The following technical elements are expressed in the BDV Reference Model:

- **Data Visualization and User Interaction**: Advanced visualization approaches for improved user experience. This technical priority is addressing the need or advanced means for visualization and user interaction capable to handle the continuously increasing complexity and size of data to support the user in exploring and understanding effectively Big Data.
- **Data Analytics**: Data analytics to improve data understanding, deep learning, and meaningfulness of data. The Data Analytics technical priority aims to progress data analytics technologies for Big Data in order to develop capabilities to turn Big Data into value, but also to make those approaches accessible to the wider public.
- **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. This technical priority is motivated by fast development and adoption of Internet of Things (IoT) technologies that is one of the key drivers of the Big Data phenomenon with the need for processing immense amounts of sensor data streams.
- **Data Protection**: Privacy and anonymization mechanisms to facilitate data protection. This is related to data management and processing as it is a strong link here, but it can also be associated with the area of Cybersecurity.
- **Data Management:** Principles and techniques for data management. This technical priority is motivated by the data explosion that is mainly triggered by the increasing amount of data sources (e.g., sensors and social data) and their complexity in structure.
- **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. This area is separately elaborated further in collaboration with the Cloud and High Performance Computing (ETP4HPC) communities.
- **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. In order to meet real-time needs, it will often be necessary to handle Big Data aspects at the edge of the system. This area is separately elaborated further in collaboration with the IoT (Alliance for Internet of Things Innovation (AIOTI)) and CPS communities.

The top elements, marked in yellow, are Big Data Types and Semantic identified, which often lead to the use of different techniques and mechanisms in the horizontal concerns and it is necessary to considered it for data analytics and data storage.

In addition, the Referent Model identifies some transversal elements on which Big Data RAs must be support:

- *Cybersecurity:* Big Data often need support to maintain security and trust beyond privacy and anonymization. The aspect of trust frequently has links to trust mechanisms such as blockchain technologies, smart contracts and various forms of encryption.
- *Communication and Connectivity:* Effective communication and connectivity mechanisms are necessary in providing support for Big Data. This area is separately further elaborated, along with various communication communities, such as the 5G community.
- *Standards:* Standardization of Big Data technology areas to facilitate data integration, sharing and interoperability.
- *Engineering and DevOps for building Big Data Value systems:* This topic will be elaborated in greater detail along with the NESSI Software and Service community.
- *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). These platforms facilitate the efficient usage of a number of the horizontal and vertical Big Data areas, most notably data management, data processing, data protection and cybersecurity.

To this end, Big Data Value Association (BDVA), euRobotics, CLAIRE, ELLIS and EurAI have launched the creation of the AI, Data and Robotics Association (ADRA) [17]. ADRA represents the private stakeholder community of the AI, Data and Robotics co-programmed Partnership in Horizon Europe. The wider purpose of objectives and activities are Big Data Value, AI, Data and Robotics update in objectives and activities.

The association objective is to boost European AI (Artificial Intelligence), Data and Robotics research, development and innovation, and to foster value creation for business, citizens and the environment.

Moreover, the aim of this association is:

- Boosting European competitiveness, societal wellbeing and environmental aspects
- Promoting the widest and best uptake of AI, Data and Robotics technologies and services for public, professional, and personal use;
- Establishing the excellence in science and business in AI, Data and Robotics.

5.4 Industrial Internet Reference Architecture (IIRA)

The Industrial Internet Reference Architecture (IIRA) [18] is a standards-based open architecture for Industrial Internet of Things (IIoT) systems. The IIRA pretends to exploit its value thanks to its broad industry applicability in enhancing common understanding, drive interoperability, mapping technologies and standardizing development. IIRA description and its representation are generic and at a high level of abstraction, to better support the requisite of broad applicability to different sectors (e.g., energy, healthcare, manufacturing, and transportation). The IIRA ”distills and abstracts” common characteristics, features and patterns from use cases defined in the Industrial Internet Consortium (IIC, the promoter committee of IIRA) and elsewhere. It will be refined continually as feedback is gathered from its application in the testbeds developed in the IIC and real applications. The IIRA is also intended to be independent from nowadays technologies to be a conceptual reference point for future deployments, independently from the specific technology and not to tie the IIoT community to existing ones, removing possible barriers for future applied research.

IIRA has two declared purposes:

1. It is the foundational framework for technical documents promoted by the IIC.
2. For the broader IoT community (academics, practitioners and users), it provides guidance and assistance in the development, documentation, communication and deployment of IIoT systems.

The IIRA deploys an architectural description of IIoT systems using ISO/IEC/IEEE 42010:2022 *Software, systems and enterprise — Architecture description* [19], presenting its concept according to different perspectives (intuitively defined as ”Business view”, ”Usage view”, ”Functional view”, and ”Implementation view”). Given the nature of task T2.2 of s-X-AIPI, concerning the modelling of a Reference Architecture independent from specific use cases and implementation, a short summary of the “Functional View” is here presented.

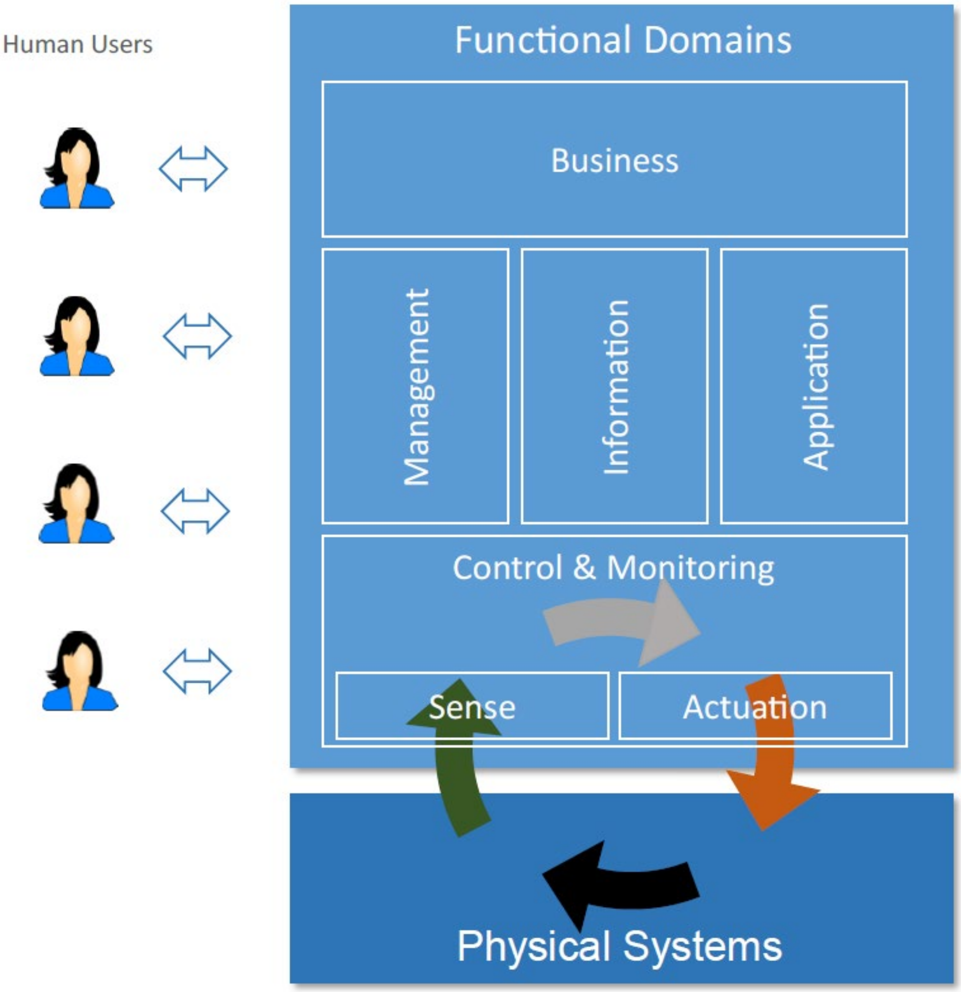


Figure 16 - Functional view of IIRA

The functional viewpoint of IIRA depicted in Figure 15 provides a five-layer description of the functions in an industrial system, their interrelation, structure, and interactions:

- Control And Monitoring domain: functions executed by industrial controllers and similar simple devices (e.g., sensing and actuation in closed-loop systems, monitoring functions and IoT communication features) as per Figure 16, where the ”IoT Gateway” block refers to those cases where a gateway is needed to connect the assets to an informative system (e.g., to connect a Profibus-communicating machine).

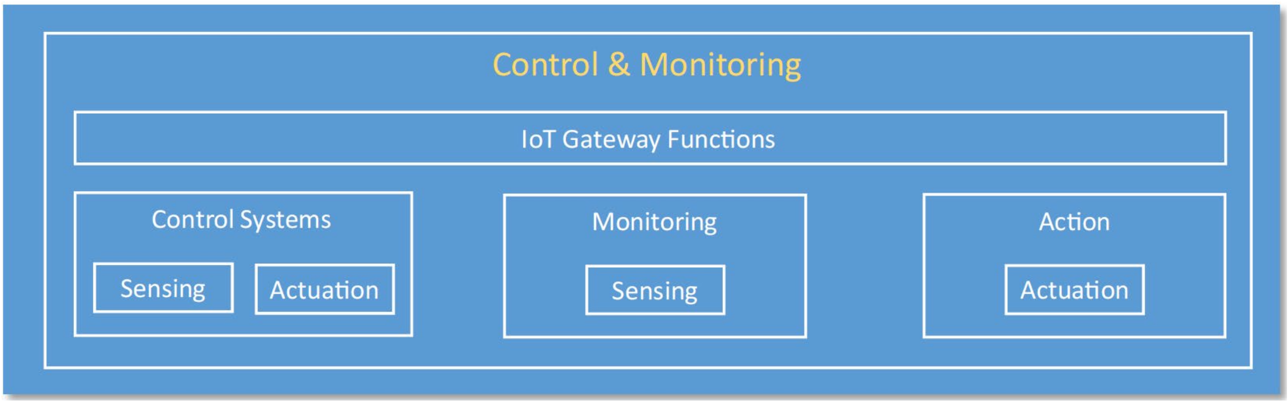


Figure 17 - Functional decomposition of Control and Monitoring domain

- System Management domain: functions that operate components throughout their life cycle (e.g., and deployment, monitoring, diagnostics of IIoT systems, as well as orchestration of different IIoT subcomponents, as depicted in Figure 17).

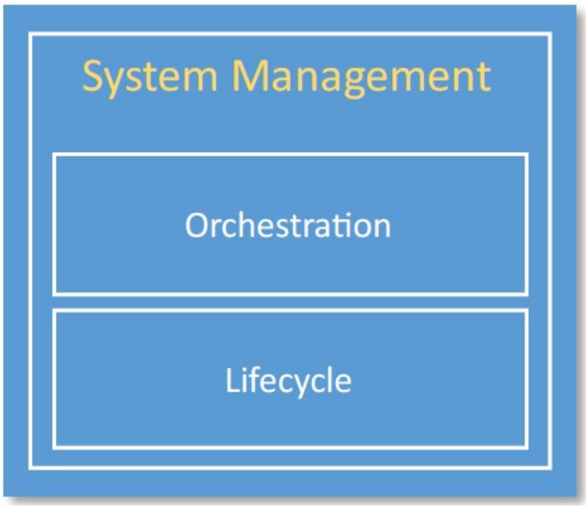


Figure 18 - Functional decomposition of System Management domain

- Information domain: functions for gathering data from the other domains (e.g., data collection and management and analytics).
- Application domain: functions implementing application logic to realize particular business functionalities (e.g., activity flows, rules and models, anyway exposed throughout User Interfaces, UIs, and Application to Program Interfaces, API).
- Business domain: functions business logics interacting with backend systems (examples of systems framed into this domain include Enterprise Resource Planning, ERP, Customer Relationship Management, CRM, Product Lifecycle Management, PLM, some functionalities of Manufacturing Execution Systems, MESs, and Human Resource Management, HRM). Figure 18 depicts the core modules and relations of the Information, Application and Business Domain.

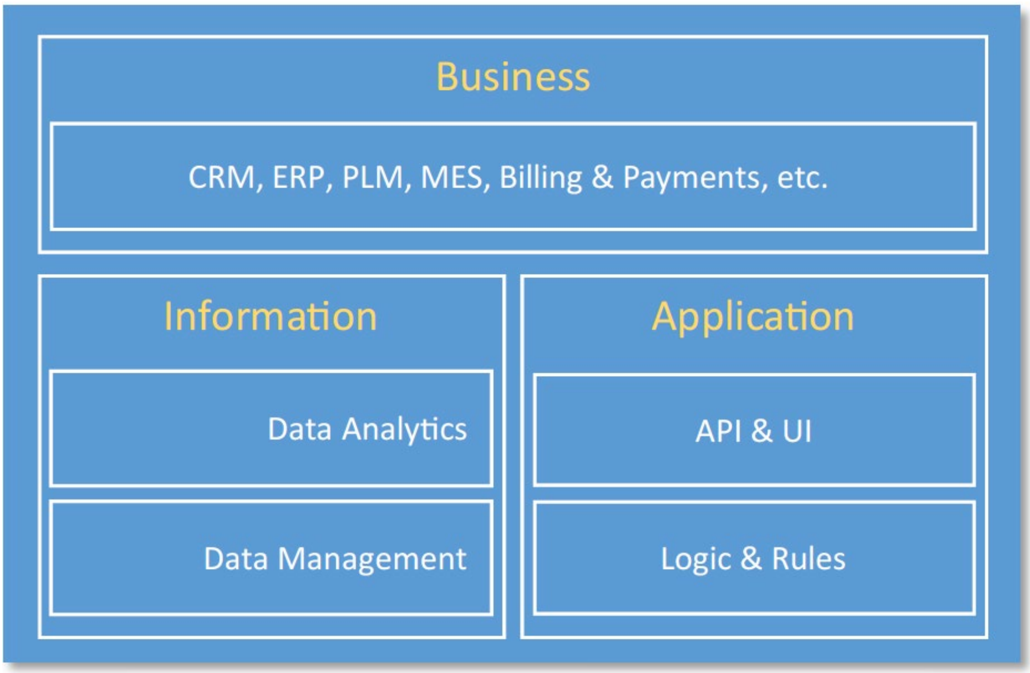


Figure 19 - Functional decomposition of Information, Application and Business domains

Figure 19 shows the alignment between IIRA and the manufacturing-focused RAMI4.0 in terms of their functional viewpoints and layers. Both models represent a taxonomy of functions associated with different layers of an enterprise, in a manner similar to the functional hierarchy model of ISA-95.

① Hardware; Software; Human Resources; Ideas; Concepts
② Sensing; Actuation; Virtualization; Modeling; Execution
③ Communication; Interfacing
④ Provisioning; Deployment; Asset Mgmt.; Monitoring; Diagnostics; Prognostics
⑤ Data Collection
⑥ Semantics; Syntaxes; Persistence; Storage; Quality Processing; Analytics; Distribution
⑦ Engines; Activity Flows; Workflows; API & UI; Service Modeling
⑧ Specification of Rules & Models
⑨ Service Orchestration; ERP; CRM; Life Cycle Mgmt.; Billing; HRM; Planning; Scheduling

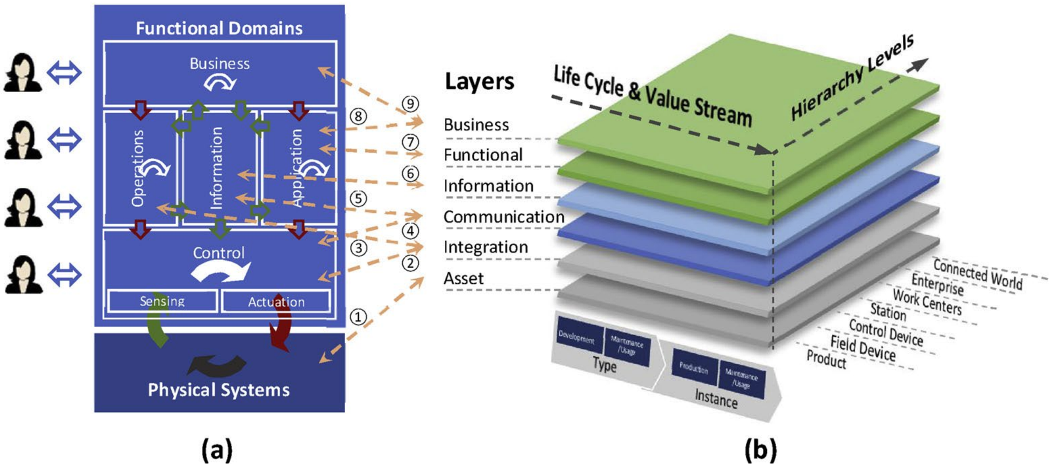


Figure 20 - Interrelations between RAMI4.0 and IIRA

Both the functional layers of RAMI4.0 and the functional viewpoint of IIRA begin from the physical assets and data acquisition and communication from the digital world to higher-level functions, data management and analytics, computational capabilities, services and business operations. CPS technologies are transforming this hierarchical representation of functions into a networked structure where the different functions can take place in a decentralized fashion and through a variety of micro-services which connects them directly to the physical assets, thanks to the IIoT communication capabilities [20] (see Figure 20).

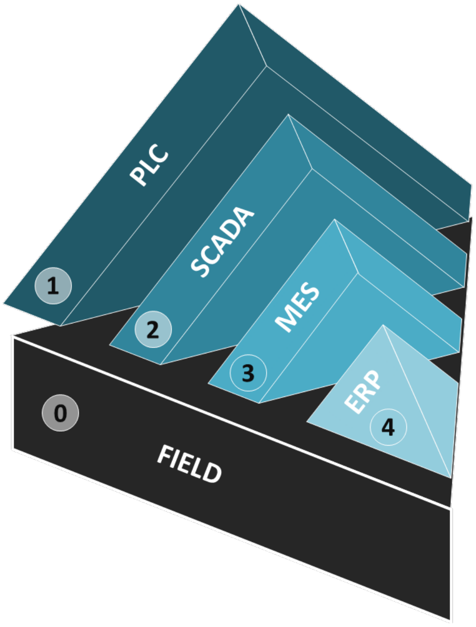


Figure 21 - "Tilted" automation pyramid [11]

5.5 International Data Space Association (IDSA)

Born as an evolution of IDS (Industrial Data Space) which itself was an initiative lead by Fraunhofer ISST, in cooperation with ATOS, T-Systems, and promoted by the German Federal Ministry of Education and Research, the International Data Space Association IDSA is characterized by the focus on generating value from data sharing, supporting information ownership and data sovereignty, via a trusted digital identity management, a fair exchange between data providers and consumers, and providing secure and reliable data sharing ecosystems and platforms. A data-driven business ecosystem might benefit from data sharing, as depicted in Figure 21, supporting for example sharing of master or event data in along the entire supply chain in Manufacturing Industry, or even, support a shared use of process data for predicted maintenance in the energy intensive industries.

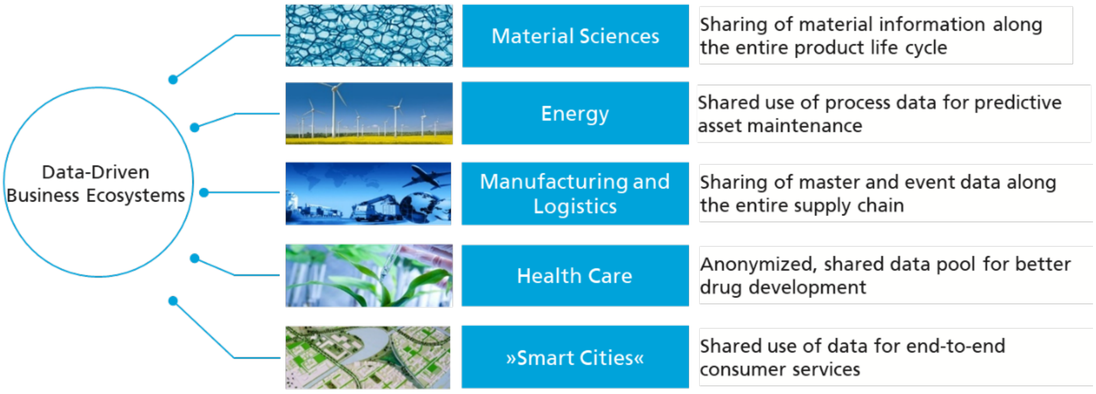


Figure 22 - Data Sharing in Business Ecosystems

To this end in 2019 IDSA suggested a reference distributed architecture that accomplishes this goal (IDS Reference Architecture Model Version 3.0) [21], now under further development: the International Data Space Reference Architecture Model (RAM 4.0) [22] aims to define the standards for implementing the data sovereignty including methods for secure data exchange and data sharing.

This has been furtherly needed due to the growing number of industrial cloud platforms, driven by technology providers, software companies, system integrators, but also existing intermediaries, and making the platform landscape very heterogeneous, with a consistent need for a standard for Data Sovereignty. Therefore, the IDS Reference Architecture Model enables links among different cloud platforms through policies and mechanisms for secure data exchange and trusted data sharing. Over the IDS Connector, industrial data clouds, individual enterprise clouds, on-premises applications and individual, connected devices can be connected to the International Data Space ecosystem, as depicted in Figure 22.

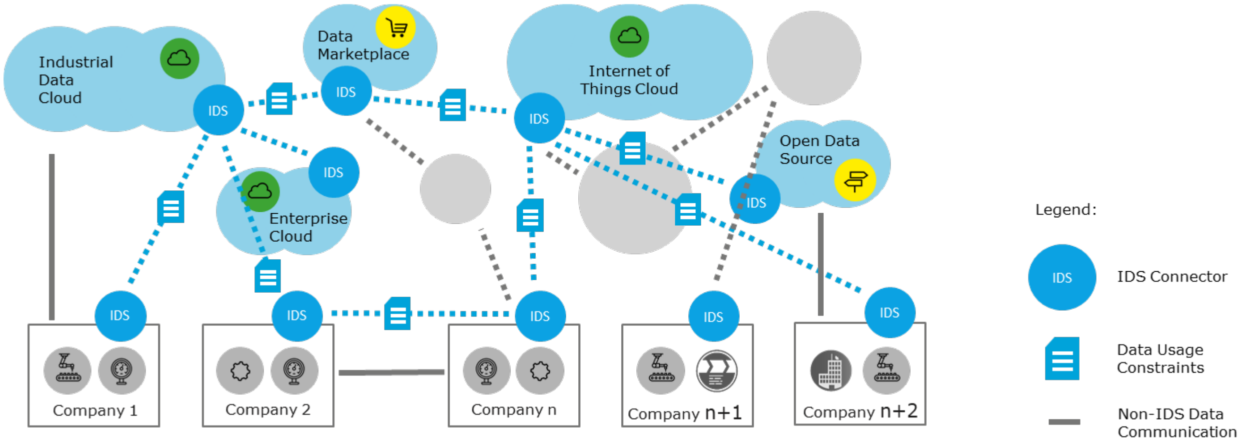


Figure 23 - Connected Industrial Platforms via the IDS Connector

The following strategic requirements are at the basis of the IDS Reference Architecture Model:

- **Trust**: enabling trust via a certified scheme in which each participant is evaluated before being granted and access the trusted business ecosystem.
- **Security and data sovereignty**: All components of the International Data Spaces rely on state-of-the-art security measures. Apart from architectural specifications, security is mainly ensured by the evaluation and certification of each technical component used in the International Data Spaces. To enable data sovereignty, a data consumer must fully accept the data owner's usage policy.
- **Ecosystem of data**: The architecture does not require central data storage capabilities. Instead, it pursues the idea of decentralization of data storage, requiring solely the description of each data sources and its usability from other stakeholders and data consumers. Furthermore, Metadata Brokers in the ecosystem provide services for real-time data search.
- **Standardized interoperability**: The International Data Spaces Connector, being a central component of the architecture, is implemented in different variants and can be acquired from different vendors. Nevertheless, each Connector is able to communicate with any other Connector (or other technical component) in the ecosystem of the International Data Space.
- **Value adding apps:** The injection of additional applications into the IDS Connectors enable the provisioning of additional services on top of data exchange processes. This includes services for data processing, data format alignment, and data exchange protocols, for example. Furthermore, data analytics services can be provided by remote execution of algorithms.
- **Data markets**: novel data-driven services are enabled by the use of data apps, fostering new business models for these services by providing clearing mechanisms and billing functions, and by creating domain-specific metadata broker solutions and marketplaces.

The above-mentioned requirements are satisfied by the IDS Reference Architecture Model (IDS-RAM) and its multiple layers - business, functional, process, information and system – and the transversal functionalities that foster security, certification and governance, as illustrated in Figure 23.

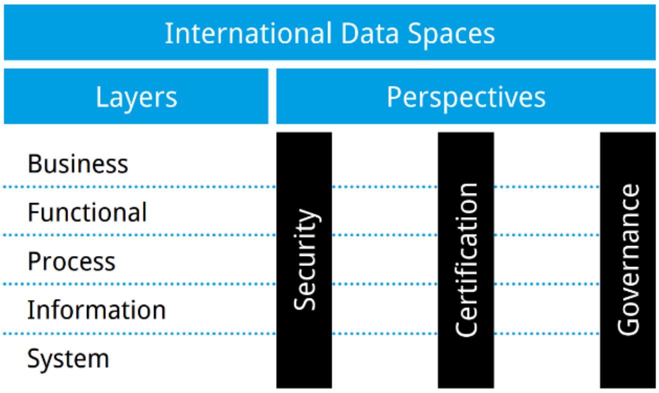


Figure 24 - IDS Reference Architecture Model

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, visualized using the BPMN notation.

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.

5.5.1 IDS Information Model

The IDS Information model [23] has been designed and defined in the context of the sub Working Group (SWG4) of the International Data Space Association. The model development is led by the Fraunhofer Institutes for Applied Information Technology FIT and Intelligent Analysis and Information Systems IAIS, and it constitutes an RDFS/OWL-ontology covering the fundamental concepts of the International Data Spaces (IDS), i.e., the types of *digital contents* that are exchanged by *participants* by means of the IDS *infrastructure components*.

The Information Model of the International Data Spaces (IDS-IM) is the central integration enabler for the semantic interoperability in any IDS ecosystem, identifying the terms and relationships to describe the IDS components, their interactions, and conditions under which data exchange and usage is possible. It represented the main principles enabling IDS communication [24].

The Information Model has been specified at three levels of digital representation, shown in Figure 24, corresponding to a conceptual meaning to the level of operational code:

- the *Conceptual Representation* of the Information Model presents a high-level overview of the main, largely invariant concepts, with no commitment to a particular technology or domain.
- the *Declarative Representation* (IDS Ontology) is the normative specification of the Information Model. It defines a fairly minimal, domain-agnostic “core model” and relies on third-party standard and custom vocabularies in order to express domain-specific facts. According to the common practice, existing domain vocabularies and standards are reused where possible, fostering acceptance and interoperability.
- the *Programmatic Representation* of the Information Model comprises a programming language data model (e.g., Java, Python, C++) shipped as a set of documented software libraries (e.g., JAR files). The Programmatic Representation provides best-effort mapping of the IDS Ontology onto native structures of a target programming language.

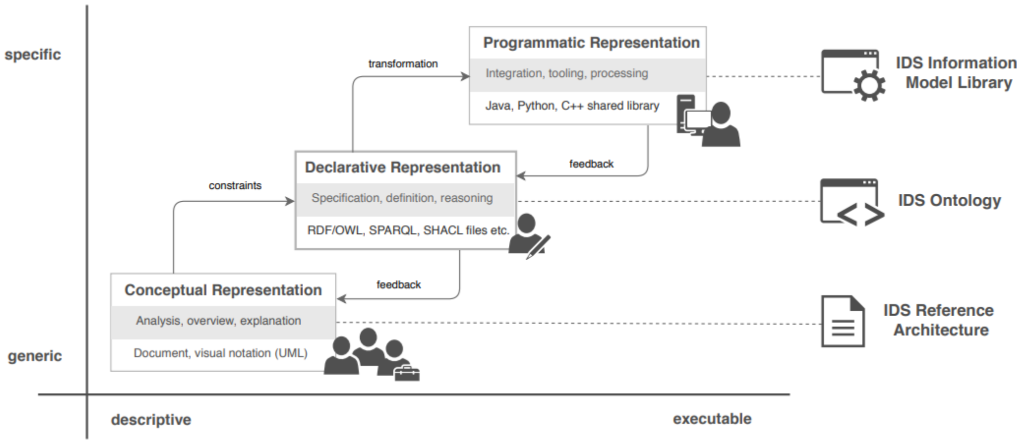


Figure 25 - Representations of the IDS Information Model

The three levels of representations provide a comprehensive overview of the IDS IM and its formal elements enabling (semi-)automated exchange of digital resources and data products within a trusted ecosystem of distributed parties, via a domain-agnostic common language. The Information Model therefore supports compatibility and interoperability, while preserving data sovereignty of Data Owners.

5.6 CAPRI

The CAPRI layered Reference Architecture has been shaped following a data-driven approach, empowered by Industrial IoT and Industrial Analytics new functions and by specific solutions to work with the process plants in real world and with the brownfield of legacy and proprietary systems, which will allow industries making strategic decisions based on data analysis and interpretation in real or near real-time. For this purpose, the core of the architecture covers a unified analytics framework offering AI-enabled services ready to integrate the advanced cognitive functionalities.

The CAPRI RA aims to support the entire data flow starting from the data collecting until the data consuming. The support for the acquisition of data from heterogeneous sources like IIoT and custom systems enables also the historicization of data generated in the IoT or industrial field on ad-hoc storage. It is guaranteed the usage of data supporting applications such as the integration of a business analytics suite based on machine learning and cognitive algorithms. The reference architecture is able to manage business analytics based on batch data and streaming data, event the edge-cloud paradigm to deploy the platform. At the same time there is the possibility to integrate security modules for the user management and adopt data sovereignty principles.

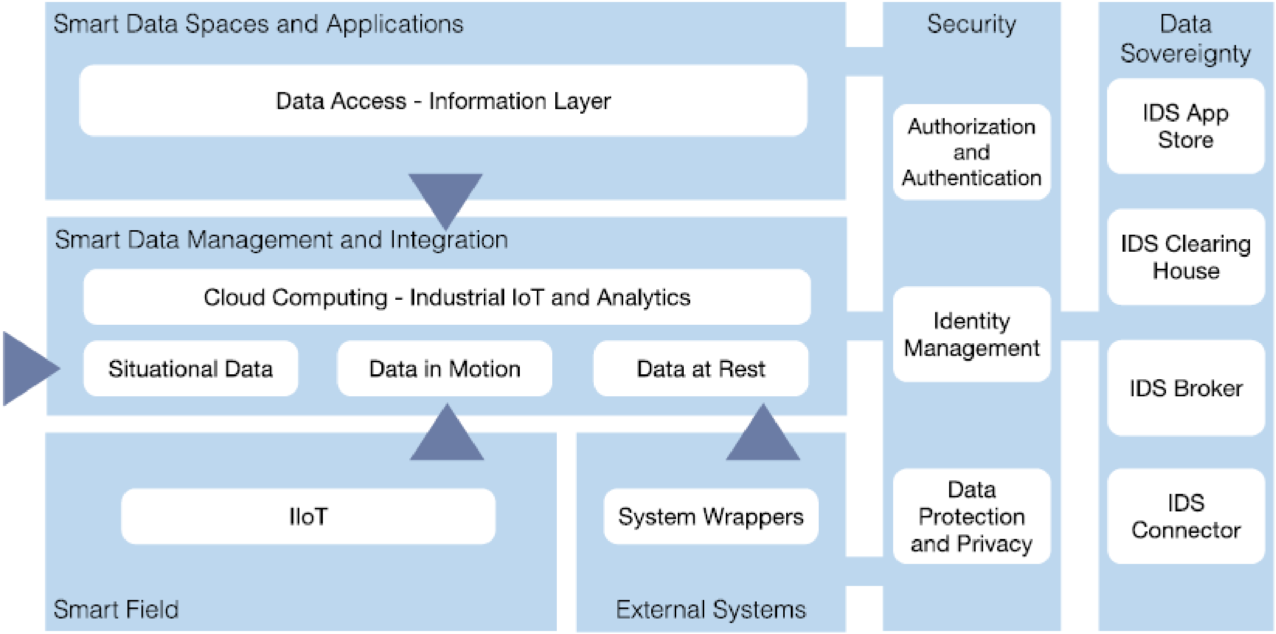


Figure 26 - CAPRI Reference Architecture

The Figure 25 depicts the CAPRI three-tier Reference Architecture, it defines several functional macro-components starting from the Smart Field containing the Industrial IoT (IIoT) physical layer, composed of machines, sensors, devices, actuators and adapters., together with the External Systems defining enterprise systems (ERPs, PLMs, customized, etc.) for the process supporting and adapters.

The Smart Field is followed by the Smart Data Management and Integration tier containing the Data Management and the Data Integration sub-modules. Regarding the Data Management, it defines information and semantic models for data representation of Data in Motion (DiM), Data at Rest (DaR) and Situational Data. Furthermore, this component is responsible for the data storage, data processing and the integration of data analytics and cognitive services.

At the top of the architecture is located the Smart Data Spaces and Applications representing the data application services to visualize and consume historical, streaming and processed data.

Two vertical layers support the adoption of Security components for the authorization and authentication of users and systems. It integrates also modules for data protection and privacy. In the same way the Data Sovereignty contains the components of the IDS ecosystem able to exchange data in a secure way guaranteeing the technological usage control and the implementation of the data sovereignty principles.

The described architecture has been conceived with modularity as a main principle: components in every layer can be combined according with a Lego-like approach, fulfilling the exposed data schema, making the architecture flexible and adaptable to the specific needs of the various application domains in process industry.

At the same time the modularity makes possible to approach a microservices design of the application that produces smaller software code, to be organizer as docker containers, so they could be run on smaller processing elements and restricted resources, as we can find in current plants, thus making easier the reuse of existing computing equipment. In this respect, the CAP Reference Architecture allows the implementation on both cloud and edge, that can be run on virtualized computing resources nearer to where multiple streams of data are created, thus addressing system latency, privacy, cost and resiliency challenges that a pure cloud computing approach cannot address, and make a big difference in process industry.

The edge implementation, see Figure 26, smoothly integrates with the cloud version, to enable data collection, storing, processing and presentation directly from the plant. Most of the short-term processing, including some data analytics, artificial intelligence and cognitive tasks could be managed at the edge, while cloud resources can be devoted to non-mission critical - massive processing of data.

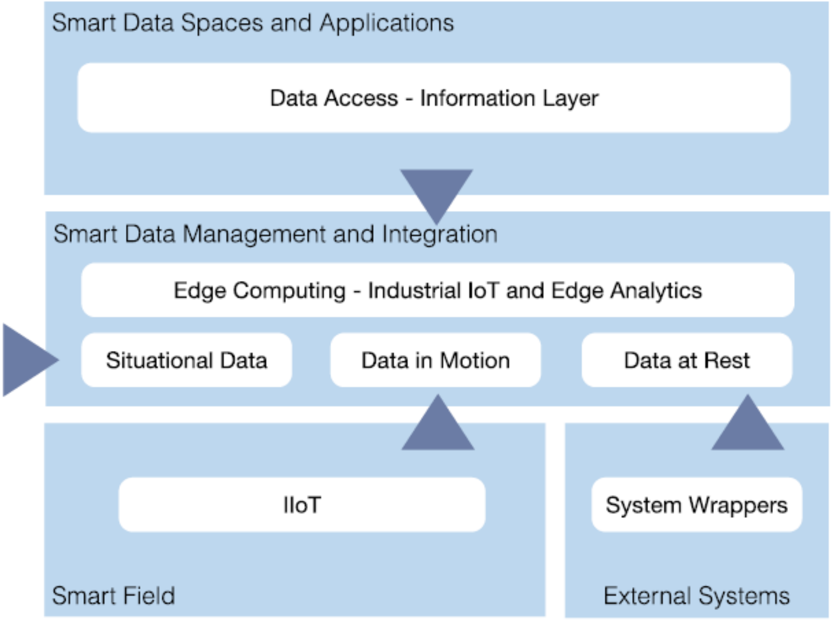


Figure 27 - CAPRI Reference Architecture for edge computing

5.7 OPEN DEI

The digital transformation strategy of the European Union identifies as a priority the creation of common data platforms based on a unified architecture and an established standard. The OPEN DEI “Aligning Reference Architectures, Open Platforms and Large-Scale Pilots in Digitising European Industry” project focuses on “Platforms and Pilots” to support the implementation of next generation digital platforms. The OPEN DEI main objective was to become an essential pillar of the development of Digitising European Industry policies, implementing four action lines:

- **Platform building:** Comparing reference architectures and open-source reference implementations, enabling a unified industrial data platform.
- **Large scale piloting:** Contributing to a digital maturity model, creating a set of assessment methods and a migration journey benchmarking tool.
- **Data ecosystem building:** Enabling an innovation and collaboration platform, forging a European network of DIHs, contributing to industrial skills catalogue and observatory.
- **Standardisation:** Conducting cross-domain surveys, performing promotion and implementation, building alliances with existing EU and standard developing organisations.

OPEN DEI [25] has the objective to harmonize and coordinate different Digital Transformation (DT) approaches under a common Reference Architecture Framework (RAF), which combines knowledge and tools, capable to foster knowledge transfer and best practices exchanges on how systems supporting DT can be architected, crossing the boundaries of specific applicative sectors. OPEN DEI proposes a conceptual model for integrated data-driven services for Digital Transformation pathways, to guide their planning, development, operation and maintenance by adopting organizations. The model is modular and comprises loosely coupled service components interconnected through a shared common data infrastructure.

The OPEN DEI project has defined the approach for designing a common Reference Architecture Framework able to describe the Cross Domain Digital Transformation. Data-driven pipelines and workflows management is nowadays crucial for data gathering, processing, and decision support. OPEN DEI has adopted the following 6C architecture to cope with this complexity, based on the following pillars (using a bottom-up reading):

- **Connection**, making data available from/to different networks, connecting systems and digital platforms, starting from the capability to make data available from/to different physical and digital assets.
- **Cyber**, modeling and in-memory based solutions to convert data into information, leveraging several information conversion mechanisms.
- **Computing**, storing and using data on the edge or on cloud.
- **Content/Context**, correlating collected data for extracting information, creating a digital space for data-information continuum, not something to push out to one side of the adopted information architecture.
- **Community**, sharing data between people and connecting stakeholders for solving collaboration needs.
- **Customization** to add value to data following each own user perspective and to match their expectations.

The above-mentioned 6C Architecture principles drive the design of the OPEN DEI RAF, developed around the main concept of Data Spaces in which data is shared (published and accessed), identifying three main different layers described in the following using a bottom-up approach:

- **Field Level Data Spaces**, it includes the Smart World Services able to collect data and support the interaction with the IoT Systems, Automation and Smart Assets (robots, machinery, and related operations) and Human Systems (manual operations, supervision, and control, etc.).
- **Edge Level Data Spaces**, it defines the typical edge operations from the data acquisition (from the logical perspective) to the data processing through the data brokering. The edge services will play a key role also for data analytics (i.e., validating and improving models for data analysis).
- **Cloud Level Data Spaces**, it includes data storage, data integration and data intelligence operations on the cloud. The cloud services will process big data, deploy algorithms, integrate different source platforms and services, provide advanced services such as AI prediction and reasoning.

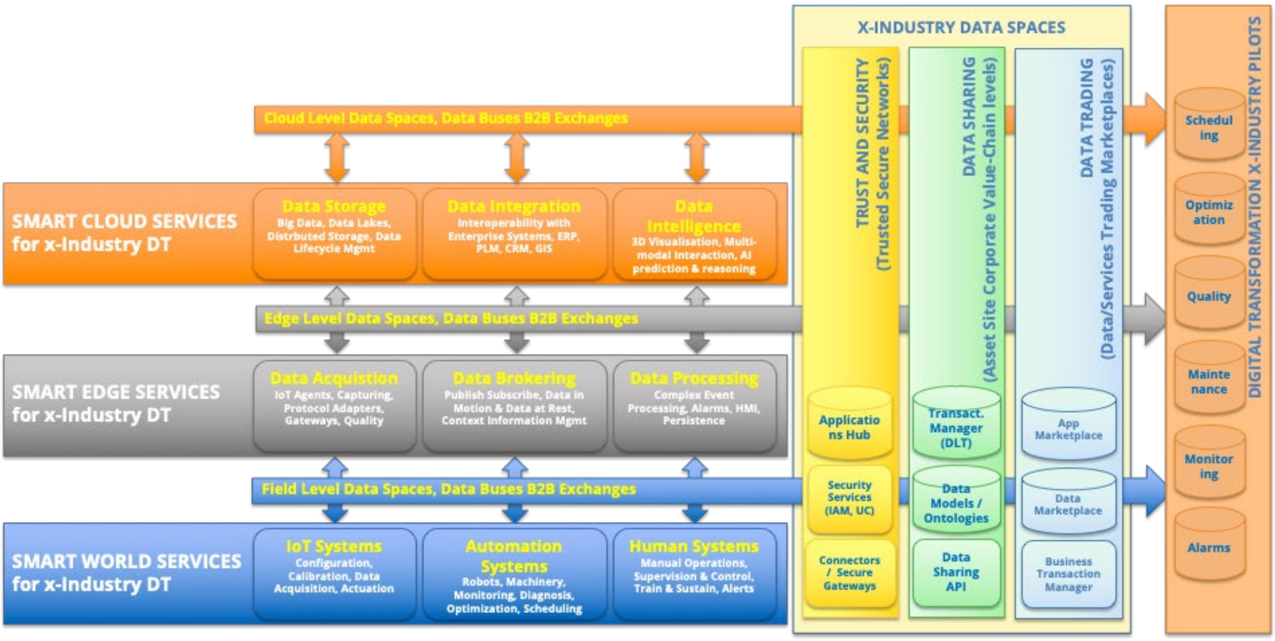


Figure 28 - OPEN DEI Reference Architecture Framework

Furthermore, all these horizontal Data Spaces tiers will feed an orthogonal dimension, named X-Industry Data Spaces, structured in:

- **Trust and Security**, including technical frameworks and infrastructures that complements the previous ones, to support trusted and secure exchange, which embraces:
    - **Applications Hub**, an infrastructure collecting the recipes required to provision applications (e.g. deployment, configuration and activation) in a manner that related data access/usage control policies can be enforced.
    - **Security Services**, a technical framework to support Identity Access Management, Usage Control and other security services.
    - **Connectors and Secure Gateways**, a technical framework for trusted connection among involved parties.
- **Data Sharing**, incorporating technical frameworks and infrastructures for an effective and auditable data sharing, which more specifically embraces:
    - Transaction Manager, a distributed ledger/blockchain infrastructure for logging selected data sharing transactions.
    - Data Models and Ontologies to leverage common standard and information representations.
    - Data Sharing API, a technical framework for effective data sharing.
- **Data Trading**, including technical frameworks and infrastructures for the trading (offering, monetization) of data, which embraces:
    - **App Marketplace**, offering applications and building blocks which can be integrated in plug&play mode to enrich existing data spaces.
    - **Data Marketplace**, offerings data resources with associated terms and conditions, including data usage/access control policies as well as pricing schemas.
    - **Business Support Functions**, enabling data/applications usage accounting as well as implementing Clearing House, Payment and Billing functions.

Finally, all the mentioned layers serve the realization of Digital Transformation X-Industry Pilots, to enable applications for supporting business scenarios from experiments. The OPEN DEI initiative and its Reference Architecture Framework fully support the Data Space conceptualization and implementation.

5.8 Other relevant initiatives

5.8.1 BEinCPPS

BEinCPPS project stated its own software architecture for manufacturing system, framing it under four different perspectives: Structural, Functional, Technical and Implementative.

The Structural Perspective is divided into two different domains and its goal is to define the basics on which to frame the entire architecture, taking into account the physical boundaries of IT systems for digitalization. It is hence divided under the Runtime System domain (including IT systems supporting operations) and Design-time System domain (including tools for the engineering of Cyber Physical Systems, CPSs). Figure 28 shows how the Runtime System domain can be framed into three further sublayers, following the fact that the module can be located on the Field (e.g., layers 0 and 1 of Figure 9), on the Factory (e.g., layers 2 and 3 of Figure 9), or on the Cloud (additional layer located on a remote, internet-connected server).

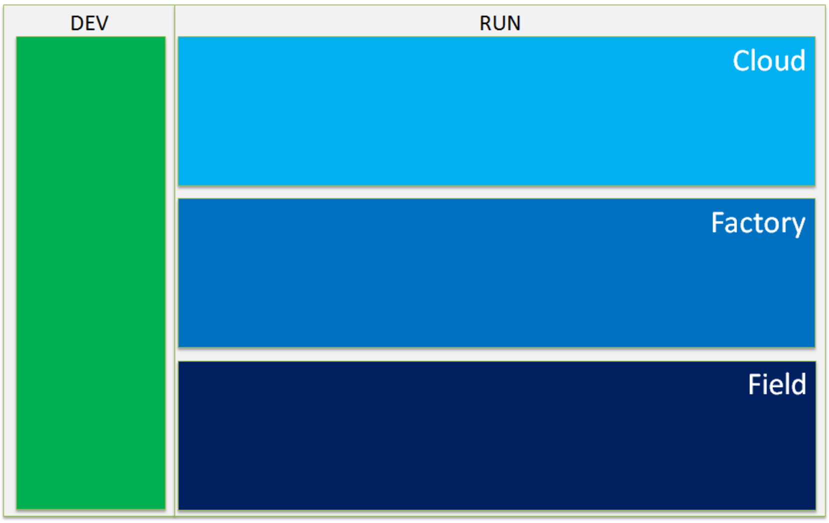


Figure 29 - Structural perspective of BEinCPPS architecture

Functional perspective is a following refining of the Structural perspective, populating it with Functional Blocks (FBs) representing functionalities and connected among each others through defined links. Hence, the Field layer links:

- Equipment & Devices: sensors, actuators, PLCs, controller boards, and machines connected to the Runtime System.
- CPPS Communication: systems providing shopfloor-level network connectivity services to equipment and devices.
- CPPS Logic: systems implementing embedded logic for monitoring and control of equipment and devices.

The Factory layer lists a unique component which is the CPPS middleware, which is in charge of providing interoperability to the data flowing through the architecture.

The Cloud layer consists of several FBs:

- Event-Data Processing: systems providing operations on data which requires notable computation resources.

- Human Computer Interaction: systems providing interactive tasks.
- Ecosystem Collaboration: systems able to support collaboration between humans and/or machines.
- Applications: systems implementing specific solutions for the end users.

On the other end, the Design-time system is divided into different FBs:

- Business Modelling: design tools for modelling organizations and business processes relevant for CPPS.
- System Modelling and Simulation: systems form modelling and simulate CPPS and simulate their behavior during useful life.
- System Engineering: tools for creating the CPPS design.

Figure 29 represents the aforementioned blocks framed into the Structural Perspective.

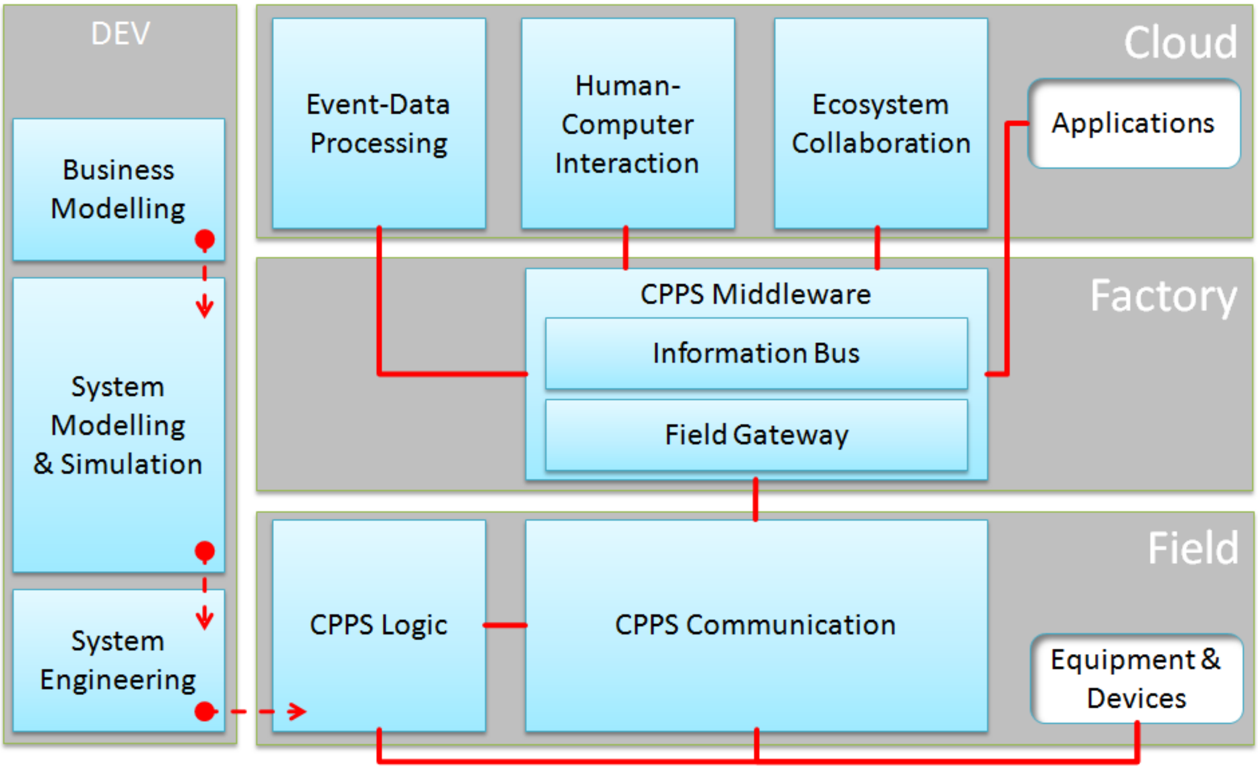


Figure 30 - Functional Perspective of BEinCPPS architecture

Technical Perspective builds over the Functional one and limits to identify a set of technical standards (either de jure or de facto) for the implementation of FBs.

Implementative Perspective is a representation which only suggests the implementation tools which can be used for the actual deployment of the project.

5.8.2 QU4LITY

Funded under the GA 825030, QU4lity was a high TRL project developing solutions for zero-defect manufacturing. In order to coordinate all the possible modules concurring to its final target, QU4LITY consortium developed a software architecture tailored to the needs of continuous quality monitoring and to efficiently connect AI modules providing core functionalities. The architecture has been named INTRA-QM and has been designed following a cloud-edge perspective.

The design of the architecture is a layered one, with highly abstracted levels hosting the “smart” applications and lower ones framing data sources and data management tools. In particular, as depicted in Figure 30 - INTRA-QM architecture, the architecture lists the following functional layers [26]:

- **Physical Elements**, intended as data sources such as sensors, IoT devices, machinery with digital interfaces and other types of CPPS systems.
- **Data Routing and Preprocessing**, the layer which is supposed to take in charge the initialization of physical elements. This layer can be seen as a connector between the physical elements and the core architecture, and its duty is the data conversion from the field raw format to an intermediate language readable by the upper layers.
- **Data bus**, the infrastructure, responsible to share data to the different recipients. In order to accomplish this task, this module lays on a middleware.
- **Processor Engine**, a programmable and configurable environment that executes data processing logic. The Processor- Engine is distributed between the cloud and edge tier of the INTRA-QM and is responsible for data processing data. At the edge tier, the Processor Engine performs low-level analytics close to the field. At the cloud tier, the Processor Engine processes data coming from different sources, thanks to the pre-processing capabilities of the layer below.
- **Device Registry**, which is a directory containing all the needed information to access and interpret specific data sources connected to the architecture.
- **Quality Management Toolkit**, used to support the integration and execution of different Machine Learning algorithms over the industrial data. The toolkit consists of two core components, respectively devoted to the procession of low-level data streams at high frequency, and of the cross-analysis of data related to different sources, operating at low frequency on aggregated data.
- **End user Dashboards,** responsible for the visualization of analytics results towards operators and managers.
- **Field Actuation,** providing interfaces from the dashboards to the field, thus enabling the human-in-the-loop concept.

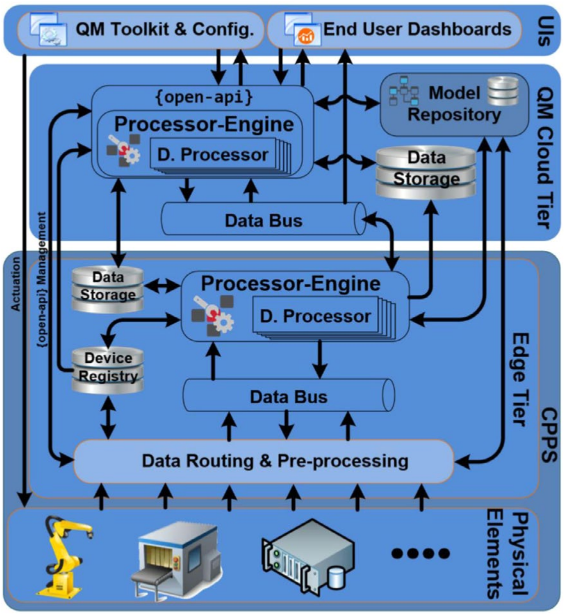


Figure 31 - INTRA-QM architecture

5.8.3 FASTEN

The EC-funded project FASTEN (GA n. 777096) tried to solve flexibility-related issues in manufacturing with a custom-designed architecture leveraging on IIoT tools and data interoperability concept. In particular, as per Figure 31, the software architecture was designed to connect field data sources (mainly sensors on production equipment) to high-level applications devoted to the monitoring and reporting of the production, to its simulation, and to the prediction of its physical outcomes (i.e., good or defected products). These application-level functionalities (namely, FASTEN Suite) were supposed to be directly interfaced with human operators, while the interfaces towards the field and the informative systems were handled through the same software architecture (embodied in the so-called FASTEN IIoT Platform).

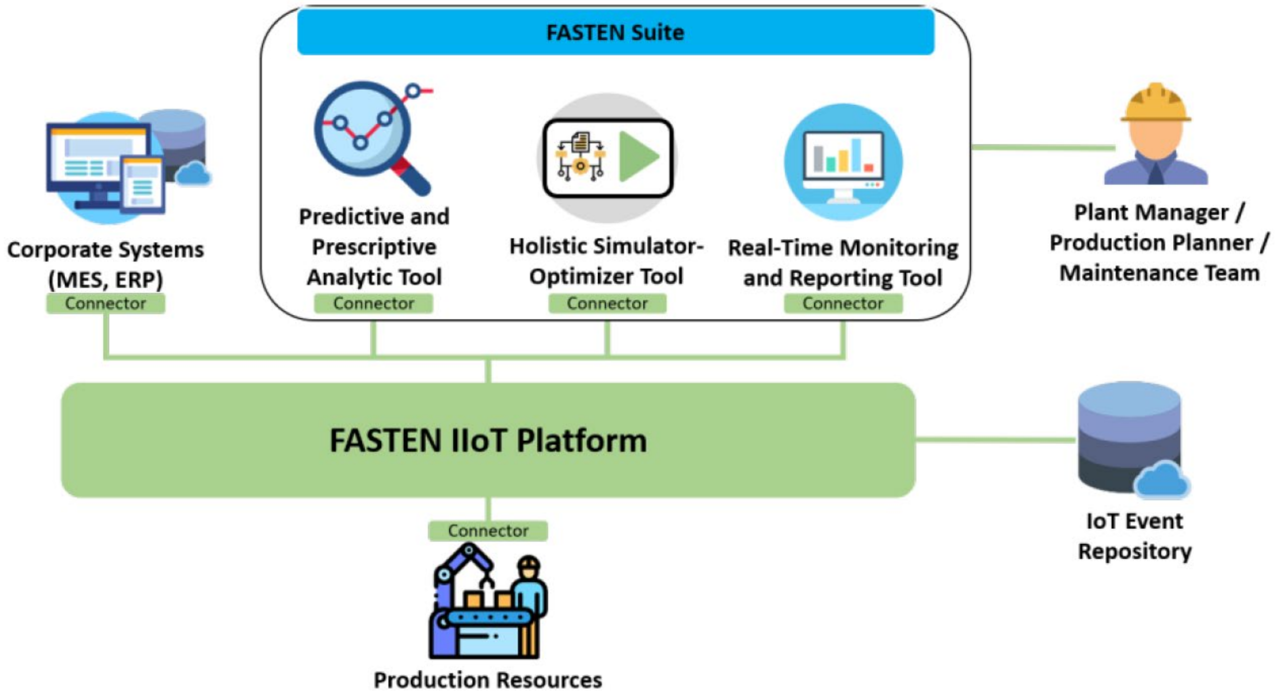


Figure 32 - FASTEN overall perspective [27]

From the point of view of the FASTEN IIoT platform, the architecture is designed as depicted in Figure 32, where the different colors represent the two possible implementation allowed by the reference architecture: in blue are highlighted the modules referring to the FIWARE catalogue, while in red are represented the modules embodied by Apache foundation related tools.

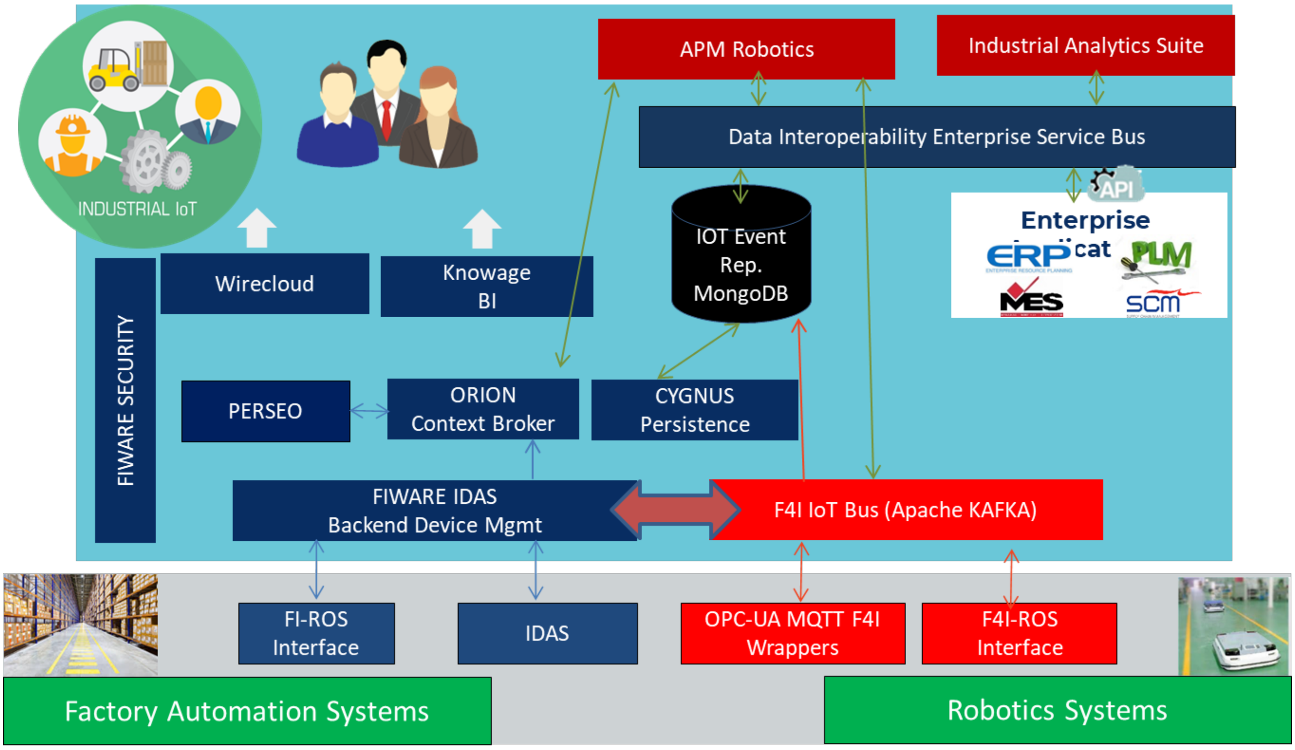


Figure 33 - FASTEN IIoT Platform [28]

In particular, the architecture can be decomposed into different conceptual layers (not aligned with the modules disposition of Figure 32):

1. Field: comparable to the layer 0 in Figure 9 - The “automation pyramid” [11]Figure 9, it represents the automated machines and systems, as well as the robotic assets supposed to be integrated in the perimeter of the FASTEN initiative.
2. Interfaces: including the blocks above the field (FI-ROS Interface [29], IDAS blocks [30], OPC-UA [31] MQTT [32], F4I Wrappers, and F4I-ROS Interface), it lists the modules in charge to route and convert the data coming from the assets addressed to the layers above. In particular, all the listed modules are in charge to filter the raw data coming from the different sources of the field, adapting them to common data formats and syntaxes, enabling the data uniformity needed for the vendor-independency and interoperability requirements.
3. Middleware: declinable in the tools of Orion Context Broker [33] (FIWARE) or Apache Kafka [34], these modules act as message-oriented boards, where the data coming from different data sources are conveyed into dedicated topics, defined and addressable through unique and pre-defined names.
4. Data handling: this layer relies on two modules of the FIWARE catalogue, where the PERSEO module is in charge of Complex Events Processing (CEP) [35], while CYGNUS connects Orion Context Broker middleware with external databases, adding some filtering capabilities.
5. IoT Event Repository: embodied by MongoDB (a non-SQL database) [36] this module is in charge to ensure data persistency for late analyses. Apache Kafka middleware grants direct a direct interface and doesn’t need additional modules.
6. Security: this layer relies on three components of the FIWARE catalogue, namely KeyRock, a token-based authentication service, Wilma [37], a proxy implementation enforcing access control, and AuthzForce [38], a module granting APIs to implement policy-based authentication decisions.
7. User Interface: this layer enables the provision of aggregated data to whom they may concern. It is indeed populated by two modules: Wirecloud [39], an environment providing mashup tools to

create dashboards, and Knowage [40], a suite providing data-based Business Intelligence and reporting functionalities.

8. Custom-based interfaces towards external tools: addressed in Figure 32 as “Data Interoperability Enterprise Service Bus” is an actual Enterprise Service Bus (ESB) [41] implementation which grants parsing and connection to external software platforms such as Manufacturing Execution Systems (MESs) [42], applications for Enterprise Resource Planning (ERP) [43] and Product Lifecycle Management (PLM) [44].
9. FASTEN proprietary tools: this last layer lists the two applications developed inside the FASTEN action, devoted to robotic assets orchestration and industrial data analysis.

FASTEN architecture can be hence considered robust and reliable, as it fulfils the typical requirements of the manufacturing domain (persistence, robustness, redundancy, and security). However, the proposed architecture relies on a classical approach, which is designed only for the transmission and management of data, but the eventual external services implementing AI functionalities need to call proper queries to recall and flag data from the static database, not implementing dedicated data pipelines.

5.8.4 Guidelines from the private sector

Apart from the aforementioned examples, which come from the public-funded domain or from associations of practitioners, notable implementations also populate the private sector of business consultancy. Constituting one of the core services of these companies, the architectures are seldom published in open source formats, but some guidelines are shared as articles targeting practitioners.

A guideline example, addressing the issue of “scaling up” the AI implementations, appeared on a specialized web magazine in 2020[2], where is given a reference architecture to implement an AI scalable platform fulfilling the requirements of scale (supporting an extended number of similar use cases), scope (extending the number of users), and speed (increase the responsiveness and the computational performances).

The proposed approach leverages on 5 main guidelines:

1. Collection of microservices: in order to better scale the implementation up, the architecture is suggested to be implemented as a collection of microservices, which can be easily deployed as a swarm of “containered” microservices (such as Docker containers or Kubernetes nodes), leveraging on the embedded tools for portability, swarming and scaling up.
2. Integration of structured approaches in the design phase: this guideline mainly aims at guaranteeing quality standards leveraging on the modules design (e.g., standardizing the workflow to scale up the production and training of algorithms or embodying traceability functionalities).
3. Enhanced data integration routes: to guarantee a faster data ingestion and to be able to leverage on big data volumes, several best practices are suggested. For example, the separation between core data storage (a highly responsive database) and the highly-entropy data (a versatile data lake) grants support from heterogeneous sources and high peak performances, as well as dedicated pipelines for AI algorithms can be used for pre-processing data on demand.
4. Event-based messaging: this feature highlights the current trend to leverage on event-based ICT tools (such as the aforementioned Apache Kafka and Orion Context Broker) to improve the performance of the platforms, avoiding at the source filtering operations on data batches.
5. Exploiting cloud components: this guideline is given as a strong suggestion in order to rely on interoperability tools natively available in cloud platforms, which, on turn allow the coexistence of open source and legacy tools by different software houses.

---

[2] https://towardsdatascience.com/how-leading-companies-scale-ai-4626189faed2

Figure 33 shows a reference architecture example based on the aforementioned principles. Following the data flow, it is remarkable how the data are pre-processed at different levels before being ingested in the data lake (3), but flow throughout modules thanks to a message bus (e.g., Apache Kafka or even SOAP-based platforms, 4), the same medium other components (such as the AI algorithms, 2) or the end-user applications (1) rely on to receive the data, properly filtered and routed by the dedicated pipelines.

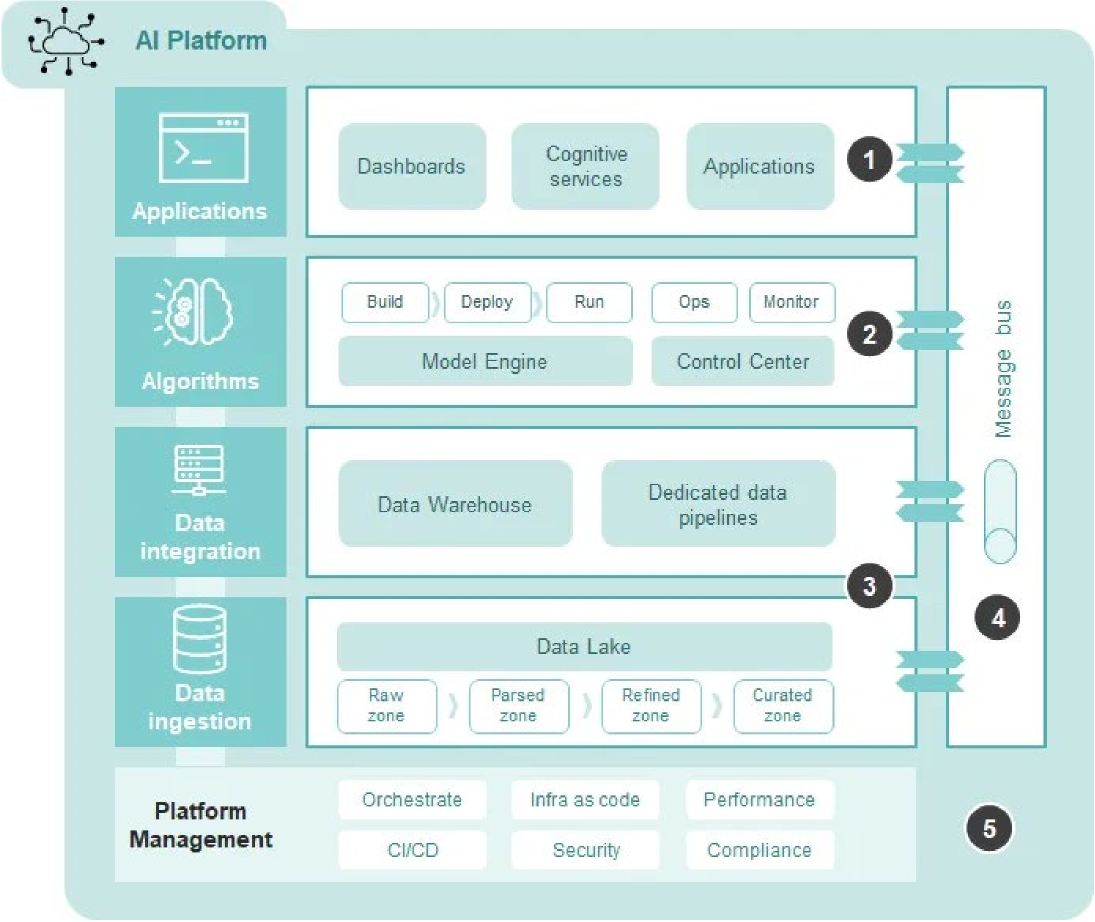


Figure 34 - Generic AI platform architecture

6 S-X-AIPI Reference Architecture and Implementation

6.1 Reference Architecture

The s-X-AIPI Reference Architecture design takes into account the main industry Reference Architectures (RAMI 4.0, BDVA, FIWARE for Industry, IIRA) and also the most relevant AI related initiatives and main background as the CAPRI, the QU4LITY and the FASTEN Reference Architectures.
The Reference Architecture supports outstanding functionalities like:

- The needs of AI-powered solution for the Process Industry;
- Human in the loop Applications;
- Autonomic Manager;
- Self-X Coordination and MAPE-K methodology;
- Interoperability with the existing industry sources.

Figure 34 depicts the s-X-AIPI Reference Architecture obtained with an iterative and co-creative process involving the domain leaders, in order to analyze domain related barriers and needs, and to integrate the outcomes of the other technical tasks of the project. The result is a "general purpose" and innovative architecture, thanks to the combination of the most relevant RAs described in Chapter 3 - Relevant Reference Architectures, able to satisfy the project requirements defined by the specific Use Cases, but applicable in different scenarios where AI pipelines are involved in.

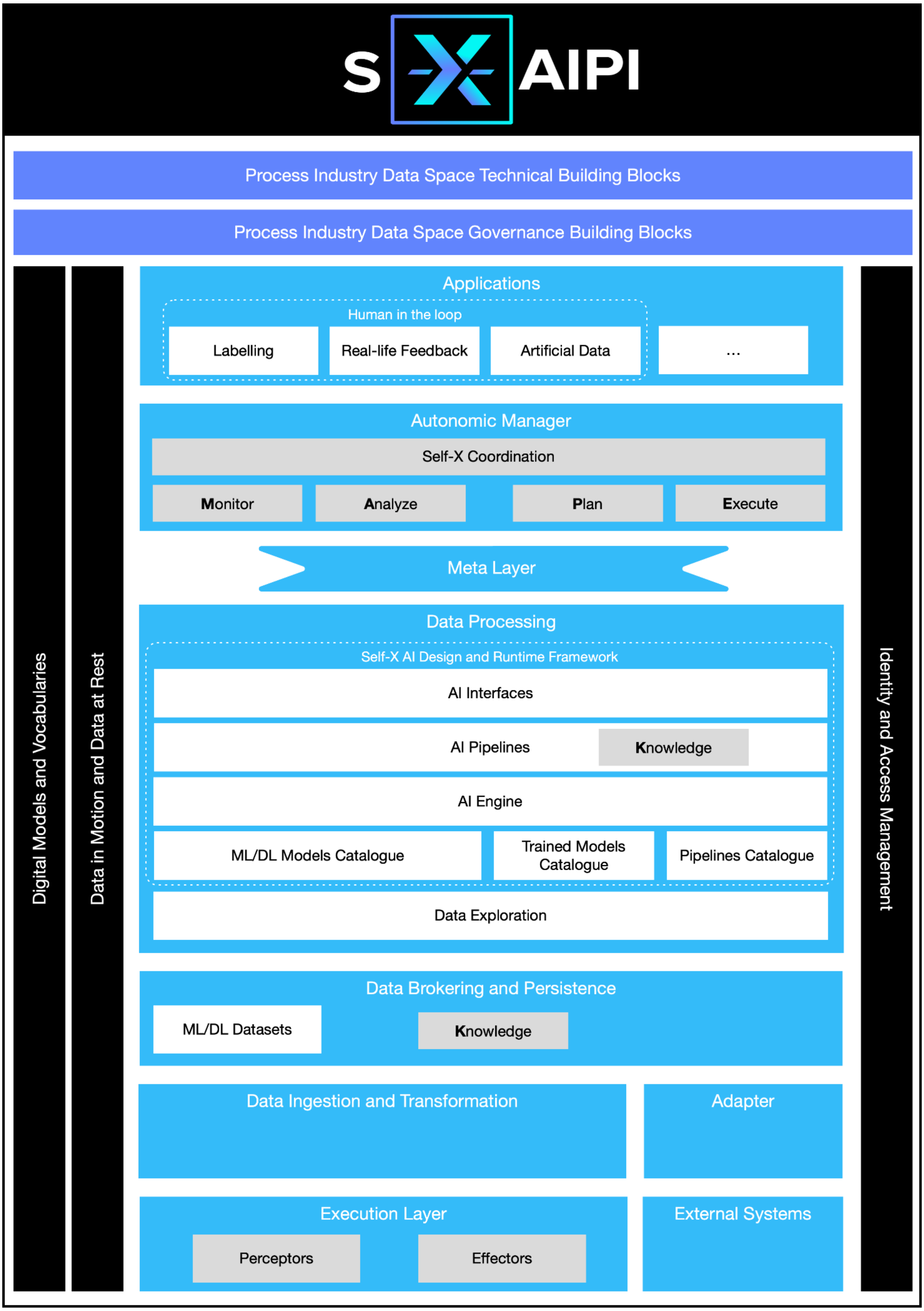


Figure 35 - s-X-AIPI Reference Architecture

The s-X-AIPI Reference Architecture is composed by the following elements:

Applications Layer

- **Human in the loop:**
    - **Labelling:** It is the process to identify datasets from raw data, building and training AI models and pipelines from scratch, provide feedback for training and tuning the models. The main purpose is to add one or more meaningful and informative label to provide a context so that the machine learning model can learn from it. This process can be supported with the human-machine interaction.
    - **Real-life Feedback:** Result methods on AI models and Pipelines, available to humans to better understand the AI model behavior, empowered with explanations, augmenting human cognition and experience with helpful insights and with interactive environments. Feedback loops are critical elements allowing to improve the performance of the models over time and ensure system sustainability.
    - **Artificial Data:** Artificially generated data produced according to real-world events, created using algorithms. Synthetic data can be deployed to validate models and to train them. They are useful to meet specific needs or conditions that may not be found in the real data like unexpected behavior, errors and unknown situations, as well as to avoid the typical issue of sensitive data sharing during the development phase.
    - Other possible applications to cover any other need
- **The Autonomic Manager** has the role of autonomous AI Data pipeline coordinator and decision maker adopting MAPE-K framework and implementing the Self-X abilities. It has the possibility also to interact with the applications layer to improve its functionalities and support the AI pipeline processing.
    - **Self-X Coordination:** Coordination of AI Data pipelines, including Self-X abilities:
        - **Self-Protection** to improve the automatic detection of possible malicious attacks on infrastructure and business data, providing the proper corrective action.
        - **Self-Configuration:** the process to automatically adapt components configuration, such as to install and setup software, add and configure new components, removing the ones old or faulty without or with a reduced human direct intervention.
        - **Self-Healing:** ability of systems to self-diagnose and self-heal without the need for operator intervention.
        - **Self-optimization:** ability of optimize hardware and software in order to allocate and use efficiently the resources, learning from experience and proactively auto-tune.
    - **MAPE-K:** The MAPE-K (Monitor-Analyze-Plan-Execute over a shared Knowledge) methodology is adopted only in the Autonomic Manager Layer to develop an autonomic and self-adaptive system, improving overall AI system and reducing human intervention. The cycle depicted in the section XXXX explains the MAPE-K behavior, assuming the availability of gathered information from the AI Data Pipelines and the execution layer (Knowledge). It can influence the future behavior of managed resources. It requires functionality to monitor the data, analyze the current status, and plan corrective actions. The core functionalities share knowledge about the managed resources and their behavior enabling the planning to respond autonomously to the needs of the system.

- **Meta Layer:** layer used to drive the communication between the Autonomic Manager and Data Processing Layers. The main feature of the meta layer is to enable Autonomic Manager for the on-the-fly problem solving, detecting problems during routing processing, find solutions to these problems, repairing and consolidating strategies to learn and to support the reuse of these solutions. On the other hand, it will complement the processing layer acting as a “repository” (push and pull) of information.

Data Processing Layer

- **AI Interfaces:** Layer including both data/results/explanation dashboarding and AI programming interfaces. The first set of tools allow users to visualize (and interact from the operator perspective) and understand model results and behavior, the second set of tools provide interfaces to build datasets, customize AI assets for training and build AI pipelines.
- **AI Pipelines:** It is the core “asset”, result of interconnected and streamlined collection of operations. They can be built from scratch by assembling pieces through dedicated Programming interfaces and executed. AI pipelines are composed of “workflows,” or interactive paths through which data moves through a machine learning platform.
- **AI Engine:** Set of tools that make possible the creation, execution and scalability of AI pipelines and models. It accesses all catalogues, acting as both an experiment platform and a design platform. It also orchestrates data flow from the input through the different AI modules of the pipeline, and finally collecting results and making them available.
- **ML/DL Catalogue:** Catalogue of the state-of-the-art ML/DL algorithms in the form of libraries, useful to create AI models from scratch. It crawls and indexes data assets stored in corporate databases and big data files, ingesting technical metadata, business descriptions and more, and automatically catalogues them.
- **Trained models Catalogue:** list of (specific domain) trained, explained and assessed models useful to build AI pipelines
- **Pipelines Catalogue:** list of (specific domain) AI applications (pipelines), provided with explanation technique, to be promptly deployed and scaled. It is a collection of reusable contents, logics, and parameters to build new solutions.
- **Data Exploration:** Data exploration layer to make data ready to be sent on the cloud platform and/or to be consumed on the edge. It is the first step of the data analysis used to explore and visualize data to uncover insights from the start or identify areas or patterns to dig into more, allows also the deeper understanding of a dataset, making it easier to navigate and use the data later.
- **Self-X AI Design and Runtime Framework:** Possibility to integrate in complex software a continuous process of incorporating self-X abilities, to increase the performance and resilience in changing uncertain environments.

Data Brokering and Persistence Layer

Collection of historical raw data from different sources (databases, data warehouses, data lakes), allowing the continuous stream of data from real-time use cases. A Data Broker aggregates information from a variety of sources transforming them in a format can be used for the upper layers. The Persistence allows to store the data in a relational or non-relational database, for the processes require this feature and also for the possibility to have historical data can be analyzed.

- **ML/DL Datasets:** List of persisted specific datasets built on the Data persistence layer, ready to be fed in AI models Pipelines. It is a collection of data that is used to train the model or related, discrete items that may be accessed individually or in combination.

Data Ingestion and Transformation Layer

Integration of the physical layer (i.e., standard and/or custom devices, external systems, etc.) transforming the information in a format/standard readable for the upper layers.

Adapter

- **Adapter:** any (protocol) transformation needed for transporting data and commands.

Execution Layer

- **Perceptors:** Collect data from physical layer.
- **Effectors:** Carry out changes through physical layer actuators

External Layer

Other: any other source or external component can produce data.

The s-X-AIPI RA described above can fully support the edge-cloud computing and it's not tailored on the specific requirements defined in [D2.1] "Scenarios and Requirements for Self-X AI adoption in Process Industry" and reported also in the following section 4.2. The needs coming from the four s-X-AIPI domains (Asphalt, Steel, Pharma and Aluminum) will validate the Reference Architecture, designed in order to be adopted in different context involving AI pipelines and autonomic computing. Furthermore, it fully supports a wider adoption in different domains (e.g., Smart City, Agrifood, Home automation, etc.), since different Data Sources in the Execution Layer do not limit its implementation.

Process Industry Data Space Technical Building Blocks

A set of technical components supporting an agile, secure, and fluid data and information flow among parties and domains. These components can be developed and deployed in several ways, based on different runtime frameworks. There are many roles within the data space including serving as fundamental building blocks that ensure data interoperability and exchange between components, such as Agents, Data Brokers, and Connectors. Other components support the creation of data value, data sovereignty and trust. Lastly, the system provides all necessary components for connecting additional systems to the data space, like the Adapters.

Process Industry Data Space Governance Building Blocks

A set of essential components to have access to relevant data and be able to share it securely with other stakeholders. The governance building blocks provide a framework for managing data in a standardized and secure manner, ensuring compliance with regulatory requirements and enabling collaboration between different parties. This framework includes components such as data sharing agreements, data ownership, access control, and privacy management, all of which are critical for building trust and promoting the exchange of data.

6.2 Mapping of the s-X-AIPI Requirements

In the following subchapters, the different functional and non-functional requirements collected in [D2.1] analysis for each use case have been mapped with the relative Reference Architecture Layer, to ensure the RA fulfillment and the subsequent integration of the self-X capabilities and the Autonomic Manager.

6.2.1 Pharma

The requirements derived in the pharma use case constitute the backbone of the desired functionality. Images from the OCT system, spectral data from the IR as well as readings from the power generator are collected in the execution layer. In the data ingestion and transformation layer, this sensor data is pre-processed and filtered.

In the data processing layer, a classification is made:

- faults of OCT images will be detected and classified;
- in the case of IR, the conversion rate will be extracted;
- concerning the power generator, the readings are verified for plausibility.

The Data Processing Layer will process and analyze relevant data in the AI pipelines, preparing the information to be subsequently exchanged with the Autonomic Manager.

The Autonomic Manager will provide extracted information to the Application Layer and external components. This will be supported by an advanced User Interface for the operator, enabling human-in-the-loop application, to control and recommend an Autonomic Manager's action or, viceversa, enabling a user action. The goal in the UI is to display critical information, enabling human intervention via an efficient data analysis.

Table 1 - s-X-AIPI mapped requirements in the Pharma domain

Requirement	F	NF	RA Layer
self-X #1 solution: "OCT image quality"			
R1. Images can be obtained from OCT 3D sensor	X		Execution Layer
R2. Images will be transferred to central lab computer via an appropriate protocol (REST-API, gRPC)	X		Data Ingestion and Transformation - Data Brokering and Persistence
R3. Images can be classified into "good" and "faulty"	X		Data Processing
R4. Faulty images can be categorized depending on the fault	X		Data Processing
R5. Sample image will be sent to UI for human in the loop	X		Application Layer
R6. Image description including fault classification will be sent to UI for human in the loop	X		Data Processing - Application Layer
R7. OCT unit must be connected to lab computer		X	External systems
R8. OCT sensor position must be determined correctly		X	External systems
self-X #2 solution: Process ML model prediction			
R1. Deviations in the OCT data can be detected	X		Data Processing
R2. Deviations in the IR data can be detected	X		Data Processing
R3. Deviations in the data from the power supply can be detected	X		Data Processing

R4. Outliers in the combined data can be detected.	X		Data Processing
R5. The quality of the surrogate model predictions is checked	X		Data Processing
R6. The surrogate model is retrained if required.	X		Data Processing
R7. Any value changes will be passed on to the UI for the user in the loop	X		Data Processing - Application Layer
R8. Retraining of model and success or failure will be reported to UI	X		Data Processing - Application Layer
R9. All lab equipment (OCT, IR, power supply) need to be connected to the lab computer and be able to communicate with appropriate protocols		X	External systems
self-X #3 solution: Control strategy			
R1. Is able to set the current in the power supply	X		Execution Layer
R2. Predicted faults can be addressed and prevented	X		Data Processing
R3. Detected faults can be reversed or prevented from getting worse	X		Data Processing - Meta Layer - Autonomic Manager
R4. Depending on the scenario, the control strategy will set the current automatically or recommend a user action	X		Data Processing - Meta Layer - Autonomic Manager
R5. Informs the UI for the user in the loop of next steps or recommended actions	X		Data Processing - Application Layer
R6. Lab power supply can process input		X	Execution Layer

6.2.2 Asphalt

The data ingestion in the Asphalt domain will be supported by a digital enabler part of the Data ingestion and transformation Layer, offering the data collected from the plant control sensors and actuators to the AI Pipelines.

For the self-X #1 solutions ”Asphalt Mix Design”, from the data ingestions layer and persistence, asphalt mix composition will be processed in the processing layer to detect deviations in key process parameters. Application and visualization layer will show and interact with the operators for production decision-making process.

For the self-X #2 solutions ”Plant Element diagnosis” in the Asphalt domain, an AI model will be integrated to determine the appropriate amount of fuel to be consumed and the electrical power, for the input of a defined quantity of the raw material to be refined. This model will be trained with the data provided by the data ingestion, together with adjustment tests at the Eiffage plant in Atalaya, modifying the variables already mentioned in order to adjust the use case considering not-controlled variables, such as the humidity of the raw material.

The above-mentioned adjustments will result in an optimum ratio final product obtained vs energy consumed.

Data Ingestion layer and Data Processing Layer will have a relevant role within self-X #3 (”Paving conditions and parameters”). Data coming, respectively, from modules attached to asphalt extending and compacting machines will have to be integrated. Data exploration will have an impact in the data processing layer, and the dashboards to propose improvement within the asphalt mix design will be part of the Application and Data visualization layers.

Regarding self-X #4 solution (”Quality Traceability at lab level”), data ingestion layer will have a major role integrating the lab data coming from tests analysis performed at EIFFAGE laboratory. Data processing layer will deal with the verification process of the results obtained at laboratory tests with the real asphalt mix design to be achieved. The application and visualization layers will provide the needed interaction for decision-making process for the corresponding users.

Within all these AI solutions to be developed the Autonomic Manager philosophy will be follow-up to minimize the participation of the data scientist in the phases of data-based AI models generation and reconfiguration.

Table 2 - s-X-AIPI mapped requirements in the Asphalt domain

Requirement	F	NF	RA Layer
self-X #1 solution: “Asphalt mix design”			
R1. Define data export procedure (Data Ingestion) from Connected Plan to CARTIF server		X	Data ingestion and transformation - Data brokering and Persistence
R2. Explore several intelligent techniques (supervised and/or unsupervised) to extract knowledge from all the production data currently register by the PLC data logger	X		Data Processing
R3. Definition of a self-X design procedure (Data Transform & Data Exploration & Model Training), based on CRISP-DM methodology and adapting as much as possible steps to the autonomic philosophy		X	Data Processing - Meta Layer - Autonomic Manager
R4. Development of a dashboard (Real-word usage) to show to the operators all the key process parameters values and/or other insights discovered in the data, thus supporting their decision-making process. Moreover, feedback from the human will be collected through this interface		X	Application Layer
self-X #2 solution: “Plant elements diagnosis” **Requirements “Plant elements diagnosis: Predictive maintenance”**			
R1. Anomaly detection (e.g. temperature, vibration, electricity consumption drift)	X		Data Processing
R2. Frequency (on-demand, when deviation detected)		X	Data Processing
R3. Horizon in future (e.g. maintenance periodicity)		X	Data Processing

R4. Trigger repairs (e.g., adjust parameters or activate recovery actions)	X		Data Processing - Meta Layer - Autonomic Manager
R5. Optimization / increase production quality	X		Data Processing - Meta Layer - Autonomic Manager
R6. Stabilization / avoid failure / enhance production continuity	X		Data Processing - Meta Layer - Autonomic Manager
R7. Final decision triggered by human via UI		X	Application Layer
self-X #2 solution: "Plant elements diagnosis" **Requirements "Plant elements diagnosis: Burner efficiency diagnosis"**			
R1. Anomaly detection (e.g. fuel consumption or temperature variation – Data exploration to evaluate the burner efficiency)	X		Data Processing
R2. Power (on-demand, when deviation detected)		X	Data Processing
R3. Horizon in future (e.g. model parameter optimization i.e. optimization of set up parameters)		X	Data Processing - Autonomic Manager
R4. Trigger repairs (e.g., adjust parameters or activate recovery actions)	X		Data Processing - Autonomic Manager
R5. Optimization / increase burner efficiency	X		Data Processing
R6. Final decision triggered by human via UI		X	Application Layer
self-X #3 solution: "Paving conditions and parameters"			
R1. Improve the current quality of the data: expand/modify the current monitoring hardware for logistics data collection		X	Execution Layer
R2. Define data export procedure (Data Ingestion) from IT ASPHALT to CARTIF server	X		Data ingestion and transformation - Data brokering and Persistence
R3. Unsupervised analysis of the data (Data Exploration) to evaluate the relationships between the production data (temperature of the material at the plant outlet) and the laying and compacting data (temperature of the material on site)	X		Data Processing
R4. Definition of the self-X design procedure (Model Training) that allows selecting the best prediction model (supervised learning) from among several possible ones, minimizing the participation of the data scientist in the process		X	Autonomic Manager
R5. Self-X adjustment of an intelligent model (Real-world usage) to predict the temperature of the asphalt upon arrival at the job site		X	Data Processing - Meta Layer - Autonomic Manager

R6. Development of a dashboard (Real-world usage) that allows the plant manager to be shown possible improvements in the asphalt mix design based on the temperature prediction. Moreover, feedback from the human will be collected through this interface	X		Application Layer
self-X #4 solution: "Quality traceability at lab level"			
R1. Define data export procedure (Data Ingestion) from HCLAB to CARTIF server based on FIWARE		X	Data ingestion and transformation - Data brokering and Persistence
R2. Verify that the results obtained in the laboratory tests complies with the asphalt mix design and comply with the standards.	X		Data Processing
R3. Develop a volumetric/mechanical properties prediction model based on gradation properties of the mixes	X		Data Processing
R4. Definition of a self-X design procedure (Model Training) that allows selecting the best algorithm or model configuration (supervised learning) to support the data analyst in future model reconfiguration tasks		X	Data Processing - Meta Layer - Autonomic Manager
R5. Development of a dashboard (Real-world usage) to show (to laboratory operators) indication of quality parameters from laboratory, its compliance with the standards and/or the prediction of their volumetric/mechanical properties	X		Application Layer

6.2.3 Steel

Steel use case will provide to the infrastructure a large set of data from which the Data Processing Layer has to identify the anomalies, update the model, monitor the production and involving the human operators for final decisions.

Contrary to the current situation, the newly developed system will support the operator with real-time insights about the scrap properties and furnace temperature and chemistry, including proposals for possible corrective control measures under the minimum operational cost.

The autonomic manager using AI data processing monitors and adapts accurate predictions of the followings:

- the scrap properties like chemical composition, especially regarding disturbing elements like copper, sulphur or phosphorus (considered as residual elements which downgrade the steel quality), and specific melting energy
- the produced steel quality in terms of temperature and composition in the EAF.

These predictions are then the basis for optimizing the performance of the EAF process in terms of material input and energy demand for achieving the required steel temperature and chemical composition.

Table 3 - s-X-AIPI mapped requirements in the Steel domain

Requirement	F	NF	RA Layer
self-X #1 solution: "Resilient high-quality raw steel production via EAF"			
R1. Anomaly detection	X		Data Processing
R2. Identification of root cause	X		Data Processing
R3. Update of scrap properties with uncertainties and trust regions	X		Data Processing / Autonomic Manager
R4. Update of process modelling tools	X		Data Processing
R5. Production planning with trust regions	X		Data Processing
R6. Frequency (every heat)		X	Data Processing
R7. Horizon (200+ heats)		X	Data Processing
R8. Final decision triggered by Human via HMI		X	Application Layer

6.2.4 Aluminum

The Data Ingestion and Transformation Layer in the Aluminum use case will collect and process data from a IDALSA proprietary LINUX application, containing all product and process-related information of the aluminum process, as well as real-time laboratory data from the aluminum mixture obtained during the melting process, with the aim of supporting operators in the aluminum recipe decision-making process. The LINUX application provides information in different data formats e.g., CSV and TXT. Hence, the proper formatting of this data will be performed to be subsequently aggregated in the Data Brokering and Persistence Layer.

In the Data Brokering and Persistence Layer, all datasets will be aggregated in proper databases focusing on raw materials, heats (process), and product information of the aluminum manufacturing. Data modelling will also be implemented to standardize and improve data usability and accessibility e.g., product passport using AAS.

The databases of the Data Brokering and Persistence Layer will feed the Data Processing Layer, containing the AI pipeline and the Catalogues of ML/DL, Trained Models, and Pipelines. The processing of data will ensure data completeness and accuracy, giving special emphasis to missing and incomplete data, to be further exploited in Data Exploration.

In addition, the latter Data Processing layer will contain the AI interfaces that will interact with the Applications Layer to include the human-in-the-loop in the IDSS (Intelligence Decision Support System), in order to support operators in the selection of the aluminum recipe. In this line, the interfaces will serve different purposes: in one hand, interfaces will show the results of the suggested recipes to operators, whereas on the other hand, they will provide capabilities to the system to obtain human feedback on those suggestions.

The Autonomic Manager and the Meta Layer in the Applications Layer will ensure the proper behavior of the IDSS by managing the coordination of the self-X abilities as well as the complete flow of information following the proposed self-X architecture.

Table 4 - s-X-AIPI mapped requirements in the Aluminum domain

Requirement	F	NF	RA Layer
self-X #1 solution: "Aluminum mix recipe"			
R1. Aluminum recipes shall result in aluminum products within the acceptable quality requirements specified in the norms	X		Data Processing
R2. The IDSS shall rely its solution considering the available materials at the plant and their characteristics		X	Data Processing / Autonomic Manager
R3. The IDSS shall work in a product basis approach upon operator's request	X		Data Processing / Application Layer
R4. The IDSS shall be able to react to changes in the plant e.g., new raw material available	X		Data Processing / Autonomic Manager
R5. The IDSS shall work with uncertainty, when possible, exploiting the use of generative approaches e.g., missing or incomplete information, not representative information from initial chemical analysis of incoming material lots.		X	Data Processing
R6. The IDSS shall be able to consider and include human feedback, incorporating new data in the re-training/optimization processes of the algorithms.		X	Application Layer / Autonomic Manager / Application Layer
R7. The solutions proposed by the AI models shall be shown to operators via user-friendly HMIs and personalized notifications according to users' preferences.	X		Application Layer

6.3 Reference Implementation

The s-X-AIPI Reference Implementation proposed in Figure 35 lists a non-exhaustive selection of open-source technologies FIWARE and Apache based, avoiding vendor lock-in, lowering costs and barriers for developers, interoperability, and user adoption. It takes into account the integrations of the necessary tools to support the AI Pipelines and their connected services.

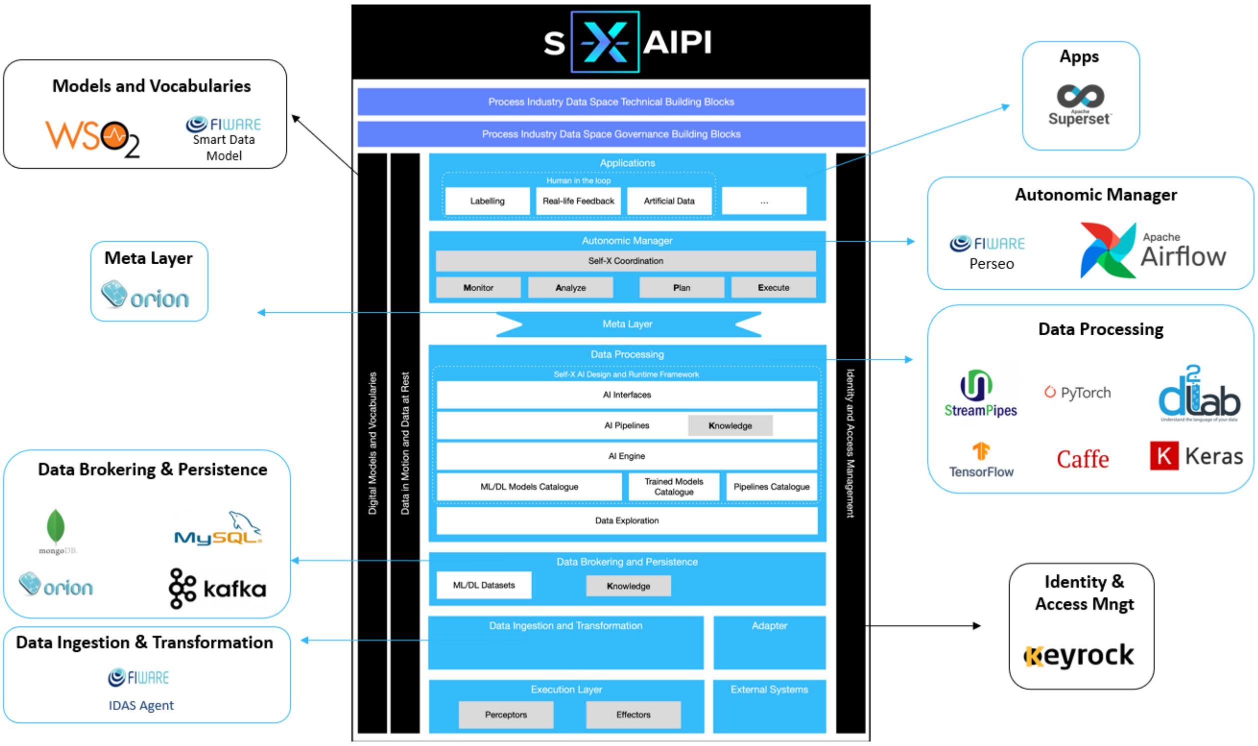


Figure 36 - s-X-AIPI Reference Implementation

Below the description of the selected technologies candidates to develop the final solution.

Models and Vocabularies

- **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.

- In the **Smart Data Models Initiative FIWARE** Foundation, TM Forum, IUDX and OASC are leading a joint collaboration to support the adoption of a reference architecture and compatible common data models that underpin a digital market of interoperable and replicable smart solutions in multiple sectors. A smart data model includes three elements: the schema, or technical representation of the model defining the technical data types and structure, the specification of a written document for human readers, and the examples of the payloads for NGSIv2 and NGSI-LD versions. All data models are public and of royalty-free nature. The licensing mode grants 3 rights to the users: free use, free modification, free sharing of the modifications.

Meta Layer

- **FIWARE ORION Context Broker** is an implementation of the Publish/Subscribe Broker Generic Enabler (GE), able to manage the entire lifecycle of context information including updates, queries, registrations, and subscriptions. It based on NGSI-LD server implementation to manage context information and its availability. This GE allows to create context elements and manage them through updates and queries, and to subscribe to context information receiving a notification when a condition is satisfied, for example in case of context change.

Data Brokering & Persistence

- **MongoDB:** is a distributed database, document-based, generic purpose for modern application and cloud. It stores documents in JSON format, supports matrix and nested objects, an advanced query language allows the user to filter data using whatever key in JSON document, having at the same time all the advantages of a relation DB like ACID transactions, the use of join in the queries and so on.

- **MySQL:** a multiplatform relational database management system (RDBMS). It is an open-source component GNU General Public Licensed developed to be aligned with the ANSI SQL and ODBC SQL standards. MySQL is a service able to manage distributed application cloud native.

- **Apache 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.

Data Ingestion & Transformation

- **IDAS Agents:** The IoT Agent component allows to connect objects to gather data or interact with them, typical IoT use case scenario. It's needed in case of connecting IoT devices/gateways to FIWARE-based ecosystems. IoT Agents translate IoT-specific protocols into the NGSI context information protocol, that is the FIWARE standard data exchange model. IoT Agent for OPC UA, IoT Agent for JSON, IoT Agent for Ultralight are some IDAS Agent in FIWARE Catalogue.

Apps

- **Apache 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.

Autonomic Manager

- **FIWARE Perseo** is an Esper-based Complex Event Processing (CEP) software designed to be fully NGSI-v2-compliant. It uses NGSI-v2 as the communication protocol for events, and thus, Perseo is able to seamless and jointly work with context brokers. The context broker tested with Perseo and officially supported is Orion Context Broker.

- **Apache 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.

Data Processing

- **Apache 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.

- **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.

- **D2LAB** is a solution as a service designed to create value by processing and analyzing a huge amount of data, used typically to detect unusual service behavior, support quality assurance, predictive maintenance and improve logistic efficiency.

Identity & Access Management

- **FIWARE KeyRock:** is a FIWARE component for Identity Management. Using KeyRock (in conjunction with other security components such as PEP Proxy and Authzforce) it is added OAuth2-based authentication and authorization security to services and applications.

List of Acronyms

AC	Autonomic Computing
AI	Artificial Intelligence
AIOTI	Alliance for Internet of Things Innovation
AAS	Asset Administration Layer
BDVA	Big Data Value Association
CAP	Cognitive Automation Platform
cPPP	contractual Public Private Partnership
DaR	Data at Rest
DiM	Data in Motion
EAF	Electric Arc Furnace
EC	European Commission
EFFRA	European Factories of the Future Research Association
HPC	High Performance Computing
IDSS	Intelligence Decision Support System
IIoT	Industrial Internet of Things
ISG CIM	Context Information Management Industry Specification Group
MAPE-K	Monitor-Analyze-Plan-Execute over a shared Knowledge
OCG AI	Operational Coordination Group on Artificial Intelligence
SRIA	Strategic Research and Innovation Agenda
WoT	Web of Trust

List of Figures

List of Tables

References

[1] P. Arcaini, E. Riccobene and P. Scandurra, ""Modeling and Analyzing MAPE-K Feedback Loops for Self-Adaptation," 2015 IEEE/ACM 10th International Symposium on Software Engineering for Adaptive and Self-Managing Systems, Florence, Italy, 2015, pp. 13-23, doi: 10.1109/SEAMS.2015.10".

[2] “FIWARE Catalogue,” [Online]. Available: https://www.fiware.org/catalogue/. [Accessed 28 March 2023].

[3] “FIWARE Marketplace,” [Online]. Available: https://www.fiware.org/marketplace/powered-by-fiware/. [Accessed 28 March 2023].

[4] “ETSI Work Item ”RGS/CIM-009v171”,” [Online]. Available: https://portal.etsi.org/webapp/workprogram/Report_WorkItem.asp?WKI_ID=66918. [Accessed 28 March 2023].

[5] European Commission, *COM/2021/206, Artificial Intelligence Act.*

[6] “Smart Data Models initiative,” [Online]. Available: https://smartdatamodels.org/. [Accessed 30 March 2023].

[7] J. Lee, B. Bagheri and H. A. Kao, “A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems,” Manuf Lett, vol. 3, pp. 18–23, Jan. 2015, , Jan. 2015, doi: 10.1016/j.mfglet.2014.12.001., p. 18–23.

[8] M. V. a. F. D. M. De Ryck, “Automated guided vehicle systems, state-of-the-art control algorithms and techniques,” Journal of Manufacturing Systems, vol. 54. Elsevier B.V., pp. 152–173, Jan. 01, 2020. doi: 10.1016/j.jmsy.2019.12.002.

[9] Scholten B, The Road to Integration: A Guide to Applying the ISA-95 Standard in Manufacturing. 2007.

[10] IEC, “IEC 62264-2: Enterprise-control system integration - Part 2: Objects and attributes for enterprise-control system integration,” 2015.

[11] W. Quadrini, E. Negri, and L. Fumagalli, “Open interfaces for connecting automated guided vehicles to a fleet management system,” in Procedia Manufacturing, Elsevier B.V., 2020, pp. 406–413. doi: 10.1016/j.promfg.2020.02.055.

[12] J. Delsing, “Local Cloud Internet of Things Automation: Technology and Business Model Features of Distributed Internet of Things Automation Solutions,” IEEE Industrial Electronics Magazine, vol. 11, no. 4, pp. 8–21, Dec. 2017, doi: 10.1109/MIE.2017.275934.

[13] S. M. e. a. Kannan, "Towards industry 4.0: Gap analysis between current automotive MES and industry standards using model-based requirement engineering," in *Proceedings - 2017 IEEE International Conference on Software Architecture Workshops, ICSAW 2017: Side Track Proceeding Institute of Electrical and Electronics Engineers Inc., Jun. 2017, pp. 29–35. doi: 10.1109/ICSAW.2017.53.*

[14] "C. , S. L. , & H. R. Manzei, Industrie 4.0 im internationalen Kontext. 2016".

[15] "F. Zezulka, P. Marcon, I. Vesely, and O. Sajdl, “Industry 4.0 – An Introduction in the phenomenon,” in IFAC-PapersOnLine, Elsevier B.V., 2016, pp. 8–12. doi: 10.1016/j.ifacol.2016.12.002".

[16] "IEC 62832-1:2020," [Online]. Available: https://webstore.iec.ch/publication/65858. [Accessed 28 March 2023].

[17] "ADRA Association," [Online]. Available: https://adr-association.eu/. [Accessed 28 March 2023].

[18] IIC. Industrial internet reference architecture (IIRA). MA, USA: Industrial Internet Consortium, Needham; 2015.

[19] "ISO/IEC/IEEE 42010:2022 Software, systems and enterprise - Architecture description" accessible: https://www.iso.org/standard/74393.html.

[20] Mohsen Moghaddam, Marissa N. Cadavid, C. Robert Kenley, Abhijit V. Deshmukh, Reference architectures for smart manufacturing: A critical review, Journal of Manufacturing Systems, Volume 49, 2018, Pages 215-225, ISSN 0278-6125, doi: 10.1016/j.jmsy.2018.10..

[21] "IDS RAM 3.0," [Online]. Available: https://internationaldataspaces.org/wp-content/uploads/IDS-Reference-Architecture-Model-3.0-2019.pdf. [Accessed 28 March 2023].

[22] "IDS RAM 4.0," [Online]. Available: https://docs.internationaldataspaces.org/knowledge-base/ids-ram-4.0. [Accessed 28 March 2023].

[23] "IDS Information Model," [Online]. Available: https://github.com/International-Data-Spaces-Association/IDS-G/blob/main/Infomodel/README.md. [Accessed 28 March 2023].

[24] "Lange, C., Langkau, J., Bader, S. (2022). The IDS Information Model: A Semantic Vocabulary for Sovereign Data Exchange. In: Otto, B., ten Hompel, M., Wrobel, S. (eds) Designing Data Spaces. Springer, Cham. doi: 10.1007/978-3-030-93975-5_7," [Online].

[25] "OPEN DEI project website," [Online]. Available: https://www.opendei.eu/resources/. [Accessed 28 March 2023].

[26] Ioannis T. Christou, Nikos Kefalakis, John K. Soldatos, Angela-Maria Despotopoulou,End-to-end industrial IoT platform for Quality 4.0 applications,Computers in Industry,Volume 137,2022,103591,ISSN 0166-3615,https://doi.org/10.1016/j.compind.2021.103591.

[27] R. Vita, N. Caldas, J. Basto, S. Alcalá and F. Diniz, An IIoT-based architecture for decision support in the aeronautic industry, MATEC Web Conf., 304 (2019), DOI: https://doi.org/10.1051/matecconf/201930404004.

[28] R. Reis, F. Diniz, L. Mizioka, P. Olivio, G. Lemos, M. Quintiães, R. Menezes, F. Amadio and N. Caldas, FASTEN: an IoT platform for manufacturing. Embraer use case, MATEC Web Conf., 233 (2018) 00009, DOI: https://doi.org/10.1051/matecconf/201823300009.

[29] R. M. A. F. L. D. P. C. F. Limosani, Connecting ROS and FIWARE: Concepts and Tutorial. In: Koubaa, A. (eds) Robot Operating System (ROS). (2019). Studies in Computational Intelligence, vol 778. Springer, Cham. https://doi.org/10.10.

[30] N. Armando, S. Sinche, A. Rodrigues, J. S. Silva and F. Boavida, "IoT Management Services: A Comparative Assessment of Popular FIWARE Agents," 2022 IEEE Ninth International Conference on Communications and Electronics (ICCE), Nha Trang, Vietnam, 2022,.

[31] Mahnke, W., Leitner, S.-H., Damm, M., OPC unified architecture (2009) OPC Unified Architecture, pp. 1-339. DOI: 10.1007/978-3-540-68899-0.

[32] M. B. Yassein, M. Q. Shatnawi, S. Aljwarneh and R. Al-Hatmi, "Internet of Things: Survey and open issues of MQTT protocol," 2017 International Conference on Engineering & MIS (ICEMIS), Monastir, Tunisia, 2017, pp. 1-6, doi: 10.1109/ICEMIS.2017.8273112..

[33] Mauro A.A. da Cruz, Joel J.P.C. Rodrigues, Pascal Lorenz, Petar Solic, Jalal Al-Muhtadi, Victor Hugo C. Albuquerque, A proposal for bridging application layer protocols to HTTP on IoT solutions, Future Generation Computer Systems, Volume 97, 2019.

[34] Shadi A. Noghabi, Kartik Paramasivam, Yi Pan, Navina Ramesh, Jon Bringhurst, Indranil Gupta, and Roy H. Campbell. 2017. Samza: stateful scalable stream processing at LinkedIn. Proc. VLDB Endow. 10, 12 (August 2017), 1634–1645. https://doi.org/10.14778/313.

[35] Y. D. a. S. R. Eugene Wu, High-performance complex event processing over streams. In Proceedings of the 2006 ACM SIGMOD international conference on Management of data (SIGMOD '06). Association for Computing Machinery, 407-418, https://doi.org/10.1145/1142473.1142520, New York, NY, USA, 2006.

[36] J. B. D. K. H. S. I. H. B. K. H. L. Seongwoon Jeong, "A data management infrastructure for bridge monitoring," ," Proc. SPIE 9435, Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2015, 94350P (3 April 2015); https://doi.org/10.1117/1d22177109.

[37] Ramli, C. D. P. K., Nielson, H. R., & Nielson, F. (2014). The logic of XACML. Science of Computer Programming, 83, 80-105. https://doi.org/10.1016/j.scico.2013.05.003.

[38] Samarati, P., de Vimercati, S.C. (2001). Access Control: Policies, Models, and Mechanisms. In: Focardi, R., Gorrieri, R. (eds) Foundations of Security Analysis and Design. FOSAD 2000. Lecture Notes in Computer Science, vol 2171. Springer, Berlin, Heidelbe.

[39] D. Lizcano, G. López, J. Soriano and J. Lloret, Implementation of end-user development success factors in mashup development environments, Computer Standards & Interfaces, Volume 47, 2016, Pages 1-18, ISSN 0920-5489, https://doi.org/10.1016/j.csi.2016.02.006.

[40] N. Leite, I. Pedrosa and J. Bernardino, "Comparative evaluation of open source business intelligence platforms for SME," 2018 13th Iberian Conference on Information Systems and Technologies (CISTI), Caceres, Spain, 2018, pp. 1-6, doi: 10.23919/CISTI.2018..

[41] M. . -T. Schmidt, B. Hutchison, P. Lambros and R. Phippen, "The Enterprise Service Bus: Making service-oriented architecture real," in IBM Systems Journal, vol. 44, no. 4, pp. 781-797, 2005, doi: 10.1147/sj.444.0781..

[42] B. Saenz de Ugarte, A. Artiba & R. Pellerin (2009) Manufacturing execution system – a literature review, Production Planning & Control, 20:6, 525-539, DOI: 10.1080/09537280902938613 B. Saenz de Ugarte, A. Artiba & R. Pellerin (2009) Manufacturing executio.

[43] Yair Wand, Ron Weber, (2002) Research Commentary: Information Systems and Conceptual Modeling—A Research Agenda. Information Systems Research 13(4):363-376..

[44] Lämmer, L., Bugow, R. (2007). PLM Services in Practice. In: Krause, FL. (eds) The Future of Product Development. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-69820-3_49.