capri

Cognitive Automation Platform
for European PRocess Industry
digital transformation

Deliverable

D2.1 Reference architecture of Cognitive Automation Platform

Deliverable Lead: ENG

Deliverable due date: 30/11/2020

Actual submission date: 17/12/2020

Version: 0.9

This project receives funding in the European Commission's Horizon 2020
Research Programme under Grant Agreement Number 870062

Document Control Page	
Title	Reference architecture of Cognitive Automation Platform
Creator	ENG
Description	the present document describes the CAPRI Reference Architecture, as functional and modular architecture that will support the IIoT, Big data and Artificial Intelligence technologies that are expected to boost the improvements and innovations in the process industry, with particular attention at the emerging cognitive technologies and solutions.
Keywords	Cognitive Automation Platform, IIRA, RAMI4.0, BDVA, IDSA, SPIRE, CPS, Artificial Intelligence (AI), IIoT, Big Data, Analytics, Data-at-rest, Data-in-motion, FIWARE4Industry, APACHE, Digital Twin
Contributors	CAR, ENG, RCPE, POLIMI, BFI, EIFF, SIDE, MSI, AIMEN, NISSA, AMS
Creation date	10/09/2020
Type	Report
Language	English
Audience	☒ Public ☐ Confidential
Review status	☐ Draft ☐ WP leader accepted ☐ Coordinator accepted
Action requested	☐ to be revised by Partners ☐ for approval by the WP leader ☐ for approval by the Project Coordinator ☐ for acknowledgement by Partners

Document History			
Ver.	**Date**	**Changes**	Author
0.1	30/09/2020	First template document with shared ToC	ENG
0.2	05/11/2020	First draft of chapter 2 and 4	ENG
0.3	16/11/2020	Review and integration of contents in chapter 4, First version of 3.1, 3.2, and 4.4, 4.6, 5.2, 5.3	POLIMI, CARTIF
0.4	23/11/2020	Use Case contribution on 5.2 and 5.3, contribution in 3.2	BFI, RCPE
0.5	30/11/2020	contribution in 3.2, revisions on 3.x, new content for 3.3 and 5.1	NISSA, ENG
0.6	10/12/2020	Improvement on 4.2, improvements on chapter 5, added 4.7, main revision on 5.2 and 5.3	POLIMI, NISSA, ENG
0.7	14/12/2020	Revisions on chapter 5, added the executive Summary, ready for final review	ENG, BFI, RCPE, CAR
0.8	15/12/2020	Some minor revisions, ready for final review	ENG
0.9	16/12/2020	Some minor revisions, addressing comments in final review	ENG

List of authors

Company	Author	Contribution
ENG	Angelo Marguglio, Gabriele De Luca	ToC Preparation, review and integration of first chapters. First draft of Chapter 2 and 4, first version of 5.1, added 3.2.1 and 3.3, chapter 6, final Executive Summary
POLIMI	Giacomo Biscardo, Silvia Razzetti	Review and integration of contents in chapter 4
CARTIF	Cristina Vega Martinez, Mireya de Diego, Daniel Gomez,	First version of 3.1, 3.2, and 4.4, 4.6, 5.2, 5.3
BFI	Christoph Nölle	Use Case contribution for 5.2 and 5.3
RCPE	Jakob Rehrl, Stephan Sacher	Use Case contribution for 5.2 and 5.3
NISSA	Nenad Stojanovic	Contribution for 3.2 and 4.7

List of reviewers

Company	Reviewer	Contribution
CARTIF	Anibal Reñones	Final review before submission

Table of Contents

This project receives funding in the European Commission's Horizon 2020
Research Programme under Grant Agreement Number 870062

Table of Figures

List of Tables

DISCLAIMER

1 Executive Summary

This deliverable, D2.1 “Reference architecture of Cognitive Automation Platform”, describes the first definition of the CAPRI Reference Architecture (CAPRI RA), as functional and modular architecture that will support the IIoT, Big data and Artificial Intelligence technologies that are expected to boost the improvements and innovations in the process industry, with particular attention at the emerging cognitive technologies and solutions.

As a first step, a general reminder highlights the objectives and scope of the project and the methodology (a conceptual model of architecture description and best practices, based on ISO/IEC/IEEE 42010 International Standard) that will be followed to describe the proposed architecture. As the same time the relationships and dependencies between the present work package, which is focused on the requirements collection and the definition of the supporting reference architecture, to be exploited in the following design and implementation of the cognitive components in the selected use cases.

A comprehensive presentation of the background of the process industry follows, how the digitalization of this segment could make processing plants and operations more agile, contributing to resource and energy efficiency, improving safety and working conditions, in line with European process industries main challenges for the future. This Digital Transformation would be the enabler for various technical innovations, innovative materials and new business models.

Then the CAPRI vision and position, with focus on the ambition to introduce the cognitive automation processes in the process industry, choosing the Asphalt, Steel and Pharma use cases as demonstrators for this.

A comprehensive description of the relevant Reference Architectures such as Industrial Internet Reference Architecture (IIRA) from IIC, RAMI 4.0 (Reference Architecture Model Industrie 4.0), IDSA RAM3.0, FIWARE4Industry, BDVA (Big Data Value Association) follows, to be considered for compliance and compatibility later in the design of the CAPRI RA.

At the same time a broad evaluation of relevant projects in SPIRE sector, with relevant contents and challenges in cognitive computing and solutions in industry, has been proposed, considering other projects like AI REGIO or AI4EU, that are exploiting AI technology for innovation.

The CAPRI layered RA has been shaped, which follows 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 CAPRI RA will cover a unified analytics framework interconnected contemplating the Data in Motion (Industrial IoT) and the Data at Rest (Industrial Analytics), offering AI-enabled data analytics services and ready to integrate the advanced cognitive functionalities. Given the compliance with well established guidelines and best practices, and covering all requirements from the case studies, the CAPRI RA can be followed in the future by process industries.

An additional check on the proposed CAPRI RA is presented on all use cases to give evidence which blocks of the RA they expect to use in the coming implementation. At the same time, the functional requirements collected in the D2.2 deliverable has been matched with the RA blocks giving evidence that all blocks in the defined RA are essential of the achievement of the projects results.

The final chapter anticipates the next steps, so how the results of the WP2 and the defined RA will be used in the WP3 and WP4 work packages to implement and integrate the use cases within the cognitive automation platform, ending up with tailoring a reference blueprints to be used by the use cases for the final demonstrators.

2 Introduction

2.1 Objectives and Scope

This is the second main outcome of the WP2, devoted to the collection of initial requirements, definition of a basic methodology and architecture and the digital transformation of process industries into cognitive process industries.

This deliverable aims to report on the initial results of tasks T2.2 (Reference Architecture for Cognitive Process Plants and Alignment with existing Cognitive SPIRE and Factories of the Future initiatives), defining the CAPRI Reference Architecture that will drive the developments in the following WP3 and WP4.

The CAPRI Reference Architecture will be aligned with state-of-the art Reference Architectures and standards; to this end, a chapter is dedicated to present an overview of the well-known Reference Architectures for the Industrial Internet-of-Things (such as IIRA, IVRA and RAMI 4.0) or data-driven architecture (like IDSA and BDVA).

As main principle, the interoperability between real (shopfloor plant) and digital world (enterprise systems) is provided by common Data Spaces at five levels sharing data through trusted databuses: data management, data protection, data processing, data analytics and data visualization Platforms. Common Data Spaces are implemented by multi-protocol data collection, multi-device and highly distributed data storage, as well as semantically enabled data models.

The Reference Implementation, to be designed and developed within the scope of WP3, will be built upon the achievements of well-known Open Source communities (such as APACHE and FIWARE) to guarantee a wider adoption and avoid lock-in.

2.2 Methodology

In CAPRI, a reference ICT architecture and semantic data model based on the ISO/IEC/IEEE 42010 International Standard will be used for representing services and IoT entities and that is transversal to many industrial sectors. This Standard, entitled as *"Systems and software engineering - Architecture Description"*, was published in 2011 as the result of a joint of the ISO and IEEE revision of the earlier IEEE Standard 1471-2000[1]. The first edition provided a conceptual model of *architecture description* and *best practices for the same*. The current edition refines the first one and adds requirements on *architecture frameworks* and *architecture description languages*.

The ISO Standard above mentioned is based upon a meta model of the terms and concepts pertaining to an architecture description intended to represent entities and their relationships. Essentially, it covers the output of an architecting process without providing any help on how to construct such a process.

Furthermore, this Standard defines an architecture framework as the conventions, principles and practices for the description of architectures established within a specific domain of an application and community of stakeholders. The fundamental goal of the architecture framework is to codify a common set of architecture practices within a community for the sake of understandability, commonality, synergy and interoperability. The Standard also defines minimum requirements on any framework that are expressed in terms of the conceptual model of Architecture Description:

---

[1] ISO/IEC/IEEE 42010: http://www.iso-architecture.org/42010/index.html

1. Information identifying the architecture framework.
2. The identification of one or more stakeholders.
3. The identification of one or more stakeholders' concerns.
4. One or more architecture viewpoints that frame those concerns.
5. Any correspondence rules.

A good architecting process should facilitate early in the overall development process the discussions about what is feasible, hard, costly, etc. In practice, it is best done concurrently with systems analysis and requirements definition activities, resulting in a set of requirements and in an architecture that meets those requirements. The inputs to this architecting process will be the system/software requirements (potentially in draft form), the system/software operational concepts, and the list of potential stakeholders. The outputs will be an architecture description that conforms to the Standard, the results from reviewing it and the architecture that it describes, and also the possibly updated requirements reflecting different architectural decisions.

CAPRI follows the following steps in this architecting process:

- Stakeholder/Concern identification. First there is a listing with all the potential stakeholders and their concerns for the system and for the architecture of the system. It will be refined and taken back to the stakeholders to be validated. As part of this process, it is desirable to extract requirements that need to be addressed in order to satisfy the needs of the stakeholders.
- Viewpoint development. The next step is to figure out how to best answer the questions listed as the stakeholders' concerns. There are two parts to this: what the answers are (*views*) and how they can be captured (*viewpoints*). Within the Standard there is an explicit separation of both of them that allows building reusable viewpoints. A well-defined set of viewpoints, reviewed by stakeholders and developers, should make it easier to capture architectural decisions.
- View development. This is basically dependent on what is required by each viewpoint. As the view is developed it is important to capture rationale for the key decisions and to include them in the architecture description. While we will not explicitly separate views from viewpoints in this document, when presenting the various viewpoints of the architecture we will keep this methodology in mind.
- View integration and evaluation. As part of the integration of the architecture each viewpoint should be re-evaluated, particularly, any rules about cross-view consistency. It has to be verified that each view correctly implements the viewpoint and that its contents fully cover the system as a whole from the perspective of that view. The architecture description should undergo some sort of evaluation of the architecture being described against stakeholders' concerns and related considerations. This is actually addressed in CAPRI via an internal review process that will iron out any inconsistences in the original viewpoint descriptions.

The methodology described here applies for the whole project, but in this report, we will focus on a Functional viewpoint that covers four levels of cognitive human-machine interaction (industrial IoT connections, smart events processing, knowledge data models and AI-based decision support), and describing consistent mapping across the existing and emerging technologies and the specified functional components in the proposed use cases. Other viewpoints will be described later in WP3 while highlighting the Reference Architecture Implementation, with a set of reference implementations both commercial and open source (WP3) for batch, continuous and hybrid process industry plants.

For documenting the architecture, it is necessary to keep a list of mandatory and optional questions that need to be answered. These concerns/topics that need to be addressed are the following:

1. Viewpoint name.
2. Viewpoint overview: a brief overview of this view and the information it presents as well as its key features or a high level view of its operations.

3. Typical stakeholders: a list of the stakeholders expected to be users of views using this viewpoint.
4. Model kinds or diagrams: identify each model kind specified by the viewpoint; alternatively provide a diagram instead of a model that outlines the components.

In addition to the previous mandatory information that should be provided, the following might be added:

1. Concerns and "anti-concerns": a list of the architecture-related concerns to be framed by this viewpoint that help to decide whether this viewpoint will be useful for a particular system of interest.
2. Correspondence rules: rules defined by this viewpoint or by its model kinds.
3. Operations on views: methods to be applied to views or to their model kinds.
4. Examples for the reader.
5. Notes: Any additional relevant information.
6. Sources: Identify the sources for this viewpoint (if any) including references.

There are important lessons taken into account and learned from the development and use of the ISO Standard. They may be summarized as falling into the following general areas:

1. **Ontology-based**. The Standard used is built upon an explicit conceptual model or ontology. To be useful, an architecture framework should be useable, that is, it must be understandable and presented in a form that can be acted upon. The first ingredient of understandability is a crisp and clear conceptual foundation and the terms reflecting it.
2. **Interest-driven**. Complex entities have a multitude of interested stakeholders each one with specific interests or concerns that, once identified, can serve as a key index into a successful architecture description.
3. **Open and extensible**. Architecture frameworks should be designed to be open and extensible as a system. One important lesson learned from architecture framework development is that defining the ontology of a given domain of interest will never be finished.
4. **Framework as a foundation**. A robust architecture framework can serve as a foundation for all aspects of architecting beyond its central role in architecture description. This means that a framework has implications for methods, processes, and tooling.
5. **Governance**. Robust conformance is a key means to achieving the usability and interoperability of architecture descriptions and should support extension, reusable model kinds and methods, as well as end-product architecture descriptions. Currently conformance is defined in terms of meta model consistency.

2.3 Relationship with other deliverables

The activities included in WP2 start from the analysis of the SPIRE context and from the specific requirements of the proposed use cases in Asphalt, Pharma and Steel domains, consolidating these requirements in the deliverable D2.2 and providing the fundamental elements to allow the design of a reference architecture in D2.1 that can meet these requirements and can be generalized to other use cases. These two deliverables form the pillars to approach the development of the modules of the Cognitive Automation Platform in WP3.

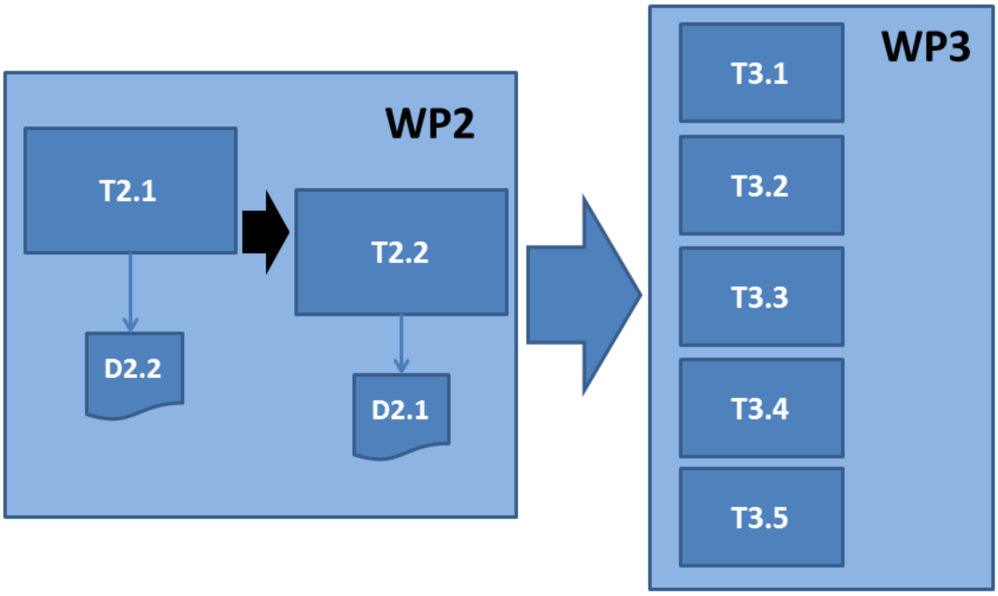


Figure 1: The relations between WP2 and WP3 tasks and deliverables.

2.4 Document Structure

D2.1 is divided in five main parts involving:

- **Introduction**: This section Identifies the tasks of the project related to the deliverable including information on objectives as well as a short description of the relationship of the current deliverable with the results of other tasks and work-packages.
- **Background and Vision**: The context of SPIRE will be described, highlighting the need for a Reference Architecture, even taking into consideration the lessons learned from other research project in the industry field. Moreover, the cognition principles will be described highlighting the fact that these can be a key element for the improvement for process industry, and so a building block to be incorporated in the CAPRI Reference Architecture.
- **Relevant Reference Architectures**: The most relevant Reference Architecture models in SPIRE will be described, and taken into consideration when approaching the CAPRI RA.
- **CAPRI Reference Architecture**: This is the core part of the document, shaping the CAPRI Reference Model, providing the solution to be adopted in CAPRI in terms of Functional Building Blocks, the required system interfaces, and the overall Functional and Modular architecture for the CAPRI project.
- **Conclusions:** This section provides summarised information on the CAPRI Reference Architecture to pave the way to the technical developments in WP3 and WP4.

3 Background and Vision

3.1 Project Background

The digitalisation of the process industries will happen horizontally and cover the entire life cycle, including R&D, plant operations, supply chain management, customer relations and integrating material flows in a circular economy and across industry sectors. It will make processing plants and operations more agile and resource and energy efficient, contribute to significant reductions of GHG emissions, orchestrate the pathway to a climate neutral economy, improve safety and working conditions and contribute to securing competitiveness and jobs in the European process industries over the next decades. The potential impact of digitalisation exceeds that of historic disruptive technological breakthroughs like the steam engine or automation (Larsson, 2019)[2].

Digital technologies will increasingly be applied in the different stages of product and process research and development. They will enable the integration of life cycle thinking and advances sustainability assessments throughout the development process. Innovation will be accelerated, leading to more efficient and much faster idea-to-market processes. One challenge is to embed customers into the R&D process. With the increasing complexity of connected industries in a circular economy, this becomes even more important. Moreover, digital technologies act as key enablers for various other innovations, such as industrial and urban symbiosis, integration of renewable energy carriers, enhancing flexibility and diversity of energy and resource inputs, innovative materials and new business models.

Materials and formulation design will be integrated with process design to achieve **digital process development and engineering**. Information about the production process and the feedstock will be included into the materials design phase and detailed characterisation method of feedstocks will be developed as input for more flexible processes.

In this context new digital enabled solutions are necessary to improve process control and operations as well as plant reliability. Significant gains in efficiency (e.g., reductions of downtime, reduced usage of raw materials and energy) can be achieved because of less rework and less waste in each production facility.

On the other hand, digital technologies can also support the design of processes using digital twins. Process design is increasingly done using faithful predictive simulation models of processes, pieces of equipment, and plants. However, several aspects still prevent the use of the full potential of model-based approach:

- Building fundamental models is a demanding process that absorbs the capacity of high-levels experts over long periods of time, so that many process elements are not described by rigorous models
- Models for the design phase are rarely used in the production phase of the life cycle of the process, and information on the behavior of the real plant and on its modifications is rarely fed back to the design model
- Fundamental data, e.g., of thermodynamic properties or kinetics is missing and expensive to generate experimentally

Technological tools are needed for digital plan operation and optimisation so that the plants can be operated in a fully energy and resource and environmentally friendly way. Developing these digital solutions will result in:

1. Demonstration of fully dynamic and model-based control of single processes
2. Decision support systems for all processes and process chains where model-based online control is not yet possible
3. Coordination and optimisation of the control of interconnected processes in a process chain
4. Development of an ecosystem of suitable digital twin of plants, processes and materials

From all points of view, it is really important the development of solutions for integrated planning, scheduling and control within plants to optimise efficiency and reliability of supply while addressing the trade-off between the maximisation of throughput and resource and energy efficiency.

---

[2] [Larsson, 2019] Larsson, Ö(2019.[AI] & Digital Platforms, Blue Institute & PII1 Insight

The success of digital solutions will depend on several factors:

- Organisations must get ready. A digital strategy is needed to integrate digital systems and platforms in the organisational structure, set the responsibilities clear and secure time of staff.
- Data integration and along the supply chain and across sectors, considering protection of IP and commercial interests and development of platforms that implement these concepts.
- Open data and resource exchange platforms that provide transparency on the quality of secondary resources, enabling a good exchange between the actors in the value chain
- The integration and interoperability of digital systems and platforms is a technical challenge and includes organisational alignment efforts, as information technology and operational technology responsibility is often spread across multiple departments.
- Data platforms to share and exchange knowledge, skill, tools, experience, and best practices provide effective communication among involved innovators/stakeholders, resulting in increasing the effectiveness and commercial benefits of their activities.
- Increase digitalisation of manufacturing industry, using the potential of digital technologies to improve innovation and productivity, leading to greener and more resilient production processes. Industrial data is increasingly collected and shared by new sensors, factory-wide communication technologies, data platforms, AI-based analytics, and more.

While digital transformation is already happening in the process industry, organising the innovations associated with digital technologies is a challenge for many. Time needed from staff to share expertise is a limiting factor in industries where margins are under pressure and there is little space for anything other than running the process.

The following section will present the most important Reference Architectures, such as IIRA (Industrial Internet Consortium Reference Architecture) , RAMI4.0 (Reference Architectural Model Industrie 4.0), FIWARE that uses a combination of technologies such as IoT, Big Data or Cloud architectures, BDVA Reference Architecture that is a reference framework made by the European BDVA (Big Data Value Association), IDSA (International Data Spaces Association), also other relevant projects in SPIRE, other projects applying AI in the industrial domain and finally a detailed explanation about CAPRI Reference Architecture.

3.1.1.1 Architecture for a cognitive process plant

In order to apply this concept to CAPRI, as explained at proposal background, different studies have reported a series of cognitive limitations analysed both at general and in each specific industrial sector related to the CAPRI pilots (i.e., minerals, chemical and steel) that can be improved with the aid of “Smart cognitive components” as IoT connection, events processing, knowledge modelling and decision support interacting with state-of-the-art automation functional architecture.

In the figure below, it is shown such interaction with traditional vertical automation levels to create an enhanced automation platform paradigm. CAPRI intends to explore this paradigm to create the envisaged Cognitive Automation Platform in order to share and exchange knowledge, tools and experience using the potential of digital technologies through creation the cognitive solutions. Following, the underlying architecture that must support this automation platform will be analysed and presented from different perspectives: CPS, Data-Buses and FIWARE-based blueprint.

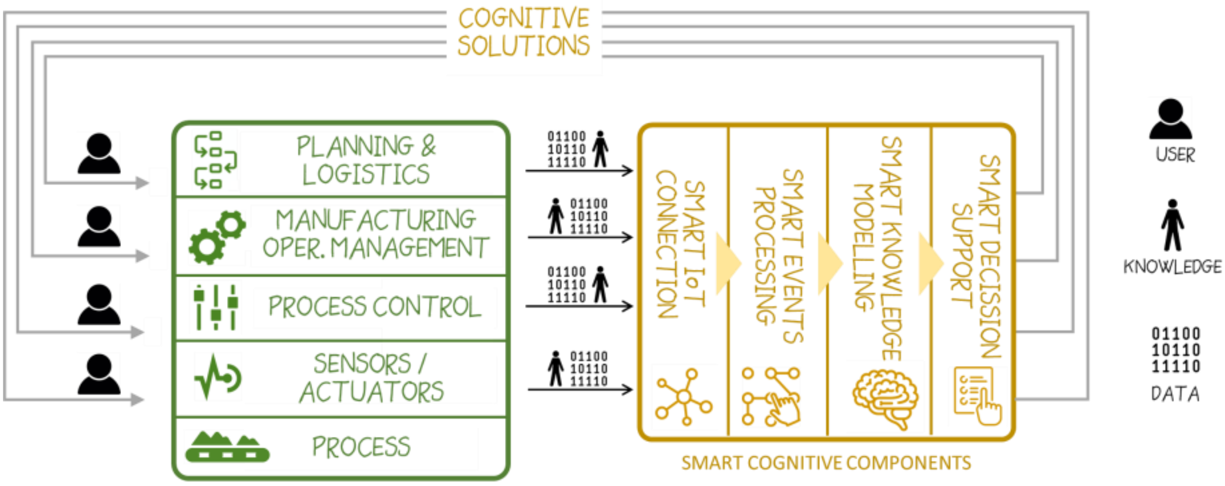


Figure 2 Automation pyramid extended with smart cognitive components

A **modular architecture** will support knowledge models, machine learning systems and different cognitive modules of planning, operation and control (noted as COGNITIVE SOLUTIONS in the previous figure). On the other hand, the architecture must be also connected with other sources of knowledge as big data flows from the factory to shop floor level (e.g., sensors, SCADA, PLC), and enterprise level (e.g., MES, ERP, PLM), as well as the information contained in the expert knowledge of workers and managers operating the plant. CAPRI consortium intends to use the data platform based on **Industrial Internet Consortium** data-driven Reference Architecture. CAPRI approach for Smart connection of data to cognitive architecture components is a data-driven evolution of the CPS automation pyramid.

At this point, it is necessary to understand **CPS** (Cyber Physical Systems), as a combination of mechatronic and digital technologies, are able to overcome the hierarchical automation pyramid and to allow cross-level interoperability organisational and technical processes, based on the interconnection between Real and Digital worlds (the physical and cyber part of our system). From a CPS perspective, the traditional pyramidal classification of Production Management Systems (Field PLC SCADA MES ERP) is substituted by continuous interoperability processes running at different levels, as very well indicated by the SCORPIUS reference model[3] in the picture below where is also shown that is possible that the levels of automation pyramid can be interconnected with each other through the use of CPSs.

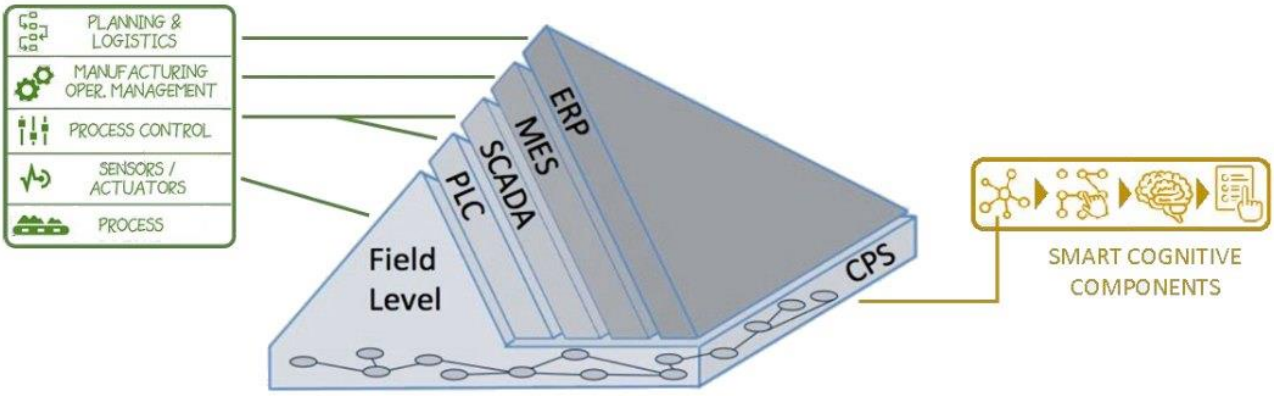


Figure 3 Product/Factory Lifecycle Automation Pyramid Vision

---

[3] https://scorpius-project.eu/. European Roadmap for Cyber-Physical Systems in Manufacturing

From the point of view of Industrial Internet Consortium Reference Architecture (IIRA[4]) interconnects industrial systems at four levels (machine, unit, site, inter-site) by several view (e.g., layered data-buses or functional viewpoints), discussing IIoT[5] challenges from an integration and interoperability perspective. IIRA and International Data Spaces Association (IDSA) have elaborated more a new interoperability paradigm, to overcome *data silos* syndrome of proprietary Enterprise Systems (PLC, SCADA, MES, ERP, …) often in the hands of different vendors, forcing them to adopt open standards in order to be interoperable.

In this perspective, IDSA is working on refining a Reference Architecture able to guarantee openness, multi-standard interoperability and data security /confidentiality over B2B and B2C relationships. Moreover adopting Cognitive approaches to extend the IIRA, on the one side, encompasses a knowledge-based interaction between humans and systems, while on the other side enables the exchange of IoT raw data (Smart IoT connection layer) at machine level, elaborated events and information (Smart Events processing layer) at unit (edge/fog) level, knowledge models and semantics (Smart Knowledge modelling layer) at site level, and generation of predictions, reports and graphics (Smart Decision support layer) at inter-site level. In summary IIRA and IDSA provide clear specifications on how to build a Digital Platform for manufacturing, while providing useful viewpoints to map existing manufacturing systems into their respective Reference Architectures. CAPRI will be able to take these references architectures as model to develop its own architecture for a cognitive process plant.

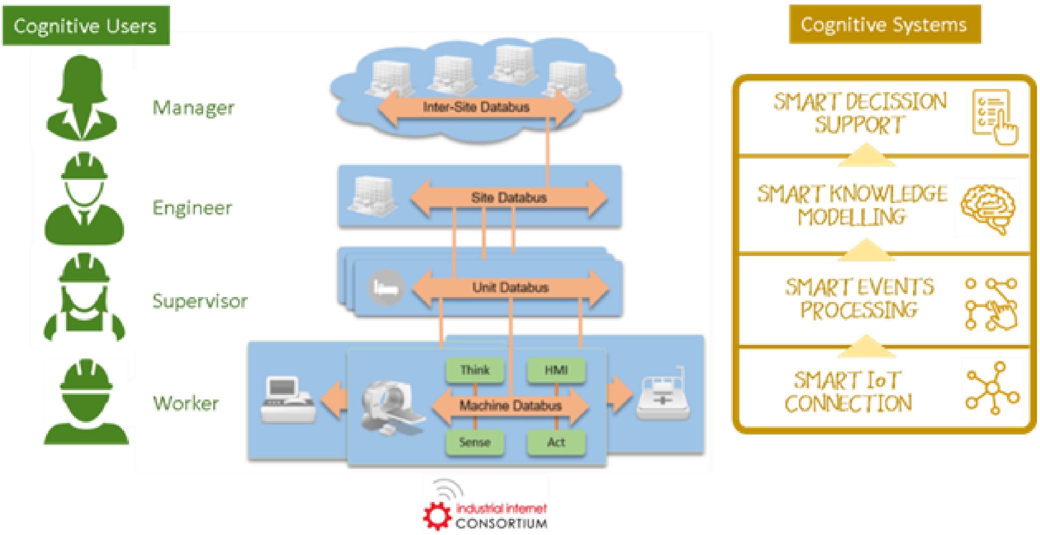


Figure 4 Layered Data-bus Cognitive Architecture

Nowadays, several initiatives are going to try to align the most adopted RAs, while different development community are actively providing their Reference Implementation. One of them is FIWARE, a worldwide community of developers providing Open Source solutions supporting Future Internet (FI) applications. This approach can be materialized in a vertical solution, covering from the field level to the applications. The figure below shows how it is possible to use FIWARE components (mentioned in the blue boxes) to build complex data- driven services based on Data in Motion (DiM) and Data at Rest (DaR). DiM represents data sources coming from the physical level of the factory (i.e., shopfloor) usually in the form of real-time feeds; DaR are more representative of factory enterprise systems (i.e., ERP, SCM, etc.), indispensable for business process and cognitive analysis.

In such a context, it is important to indicate a dedicated endeavour is being carried out to deliver a FIWARE for Industry set of components able to realize vertical manufacturing platform by reusing

---

[4] https://www.iiconsortium.org/IIRA.htm

[5] Industrial Internet of Things

several FIWARE Open Source enablers such as other components delivered by other Open Source community (e.g., within the Apache Foundation).

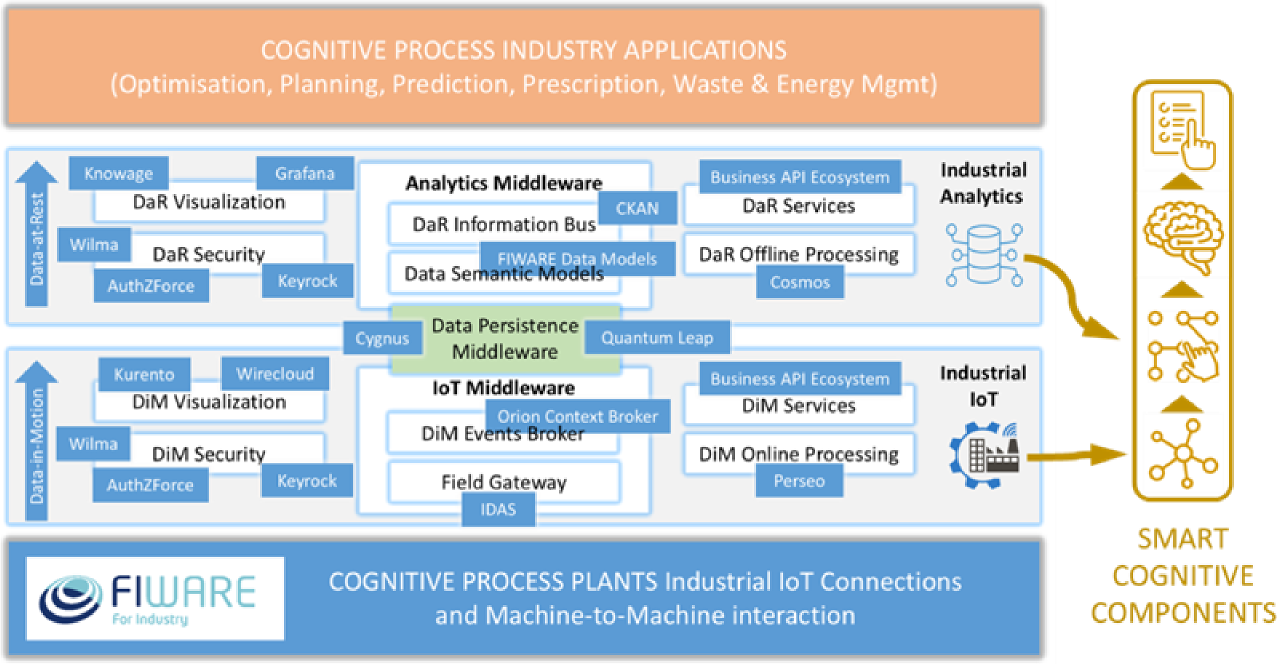


Figure 5 Cognitive Process Plants Blueprint

3.2 CAPRI Vision and Position

Process industry plants must face energy and environmental efficiency issues due to the fact that many times, raw materials, energy prices and anti-pollution policies are important constraints that must be dealt with, forcing them to adapt and adjust their process no matter their complexity.

Process industries must also face different challenges, being one of the most important the digitisation. Digitisation is usually related to the adoption of some new and promising technologies like cloud computing, the Industrial Internet of Things (IIoT) or machine learning. All of them promise to solve in an easy and cheap way problems that have affected process industries for years—creating new opportunities along the way. Smart sensors and improved data availability will enable new technical solutions and business models, and any company that does not leverage these new opportunities will have a difficult time competing with those that adopt them faster. Achieving digitalization does not require implementation of all of these new technologies; even old-style implementation of manufacturing execution systems (MES) or manufacturing operations management (MOM) systems is enabling the digital initiative. In fact, in some ways, adopting an efficient MES/MOS is the real enabler of a digital transformation and that all the other technologies are just new tools to implement simple or more sophisticated MES functionalities.

Implementing an MES is basically implementing an IT system for operations. It supports with a computerized approach the most important functions and processes that characterize the life of a specific plant. It essentially digitizes the production process, collecting, transferring and manipulating data in a digital way in order to provide information to the right people, at the right time in the right place. It must be implemented in the most functional and cost-effective technology available: smart sensors to collect data from sources that would not be available otherwise, cloud to store data and make it easily available without self-growing a complex and risky IT infrastructure, machine learning or artificial intelligence to enhance data and understand its complex correlations to create usable information, and analytics and visualization to present the information in a format that is more intuitive and effective for humans to consume.

Thanks to an efficient MES usually process industries, are able, through an advanced planning and scheduling system (APS) to generate an optimized production plan, transforming customer orders into specific production orders. The function of production performs a detailed scheduling of the production orders with allocation of necessary capacities and resources in production aggregates, energy and material amounts (usually using the mentioned MES platform).At control function, plant operators, with their specific skills and abilities, learnt through years of expertise and trial-and-error interaction, perform the appropriate fine-tuning of the control of specific unitary processes, interacting with specific HMI, SCADA and control systems and many times through direct intervention in the process.

In the traditional automation pyramid model, process industries uses local (level ‐ 2) control systems, usually based on dedicated PLC and in an upper level, SCADA systems, where they control the state of the process and send orders to the elements and actuators to perform their corresponding actions dedicated to single processes but with insufficient integration to the upper level (level ‐ 3) organizational and production management system with zones identified by similar security and communication speed requirements. Usually, PLC output variables can be obtained every very few seconds (or milliseconds) and in the following level the information does not require in many cases to be stored every few seconds, but in minutes, hours, shifts or even days.

This is usually the level of the MES systems, which includes information on productive operations, logistics, quality and safety. Any optimization at local level, however, does not imply the global/holistic optimization of the process plant. Many times, in process industries, communications between levels are precarious and the information can go through the pyramid level in ascending or descending order.

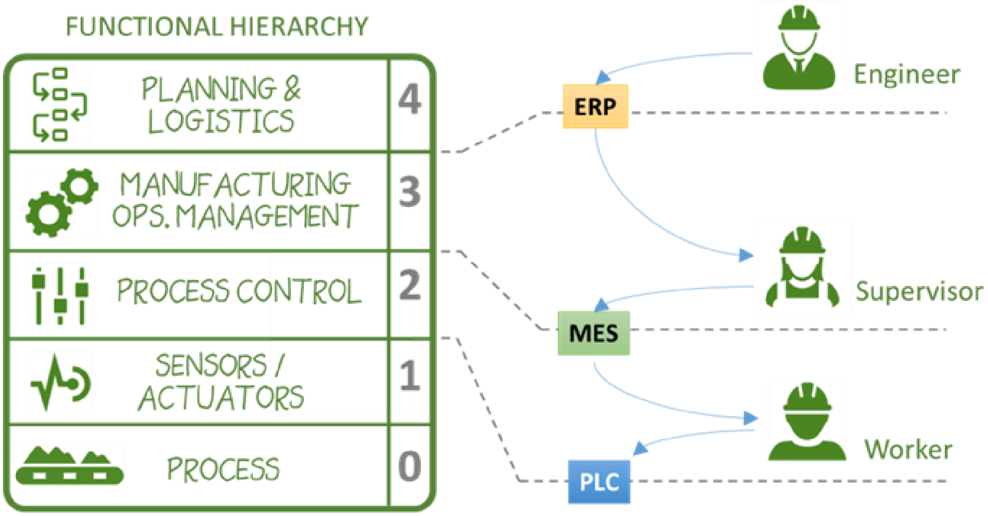


Figure 6 Classic elements of the automation pyramid in process industries

In all process technical interfaces, current static integration of hierarchical technological components (ERP, APS, MES, PLC, DCS, HMI, SCADAs, sensors and actuators) is not able to guarantee the agility and flexibility required nowadays by Industrial IoT and next generation control systems.

Moreover, in contrast with extreme automation and control, in process industries, **human factor is still essential** as a highly flexible asset for achieving a fast response against **rapid changes of factories' surrounding environment** like a sudden fault of an actuator or an unexpected change in a quality parameter of raw material.

Current digitisation approaches provide capabilities for analysis that in many cases give **insights about usual behaviour of different plant subsystems** as well as model-based monitoring and prediction of the respective processes.

Recent information and communication technologies (Industry 4.0 technologies, IIoT, Machine Learning, etc.) applied to automated production are not able by themselves to **adapt behaviours to unpredictable situations**, to take unforeseeable recovery actions, to learn and gain from experience, especially when communicating with other systems. There is the need to augment processes with **cognition capabilities** (as existing in human intelligence) to enable an efficient automatization when unpredictable behaviour happens**. This includes the detection of unpredicted situations, finding appropriate solutions, detect variations in all process chain, understand those variations and their impact and consequently, find the optimal reaction or behaviour to be applied**.

When it comes to decision-making under tough circumstances in industrial environments, managers and process operators need to deal with obstacles and constraints to rational decisions like confirmation bias (to support one's beliefs), loss aversion (to prefer avoiding losses than equivalent gains) or the illusion of control (overestimation of one's control ability) to name a few of the many cognitive biases. Such biases help address main four problems:

- **Information overload**: with too much available information at the plant, there is no choice but to filter almost all of it out. Human brain uses simple tricks to pick out the bits of information most likely to be useful.
- **Lack of meaning**: to cope with inherent industrial environment complexity, and seeing a tiny sliver of relevant information, we need to make sense to survive. Once the reduced stream of information comes in, human brain connects the dots, fill in the gaps with what we already think we know, and update mental models of the corresponding processes.
- **The need to act fast**: as every piece of new information comes in, we need to (1) do our best to assess how to affect the situation, (2) apply it to decisions, (3) simulate the future to predict what might happen next, and otherwise (4) act on our new insight.
- **How to know what needs to be remembered for later**: With too much information in the factories, we can only afford to keep around the bits that are useful in the future. Human brain makes constant bets and trade-offs around what needs to remember and forget.

A cognitive system needs to provide support for human-driven solution to such problems in process plants, represented by three sectors (minerals, steel and chemical).

Cognitive biases - Problems	MINERALS - ASPHALT	CHEMICALS - PHARMA	STEEL - BILLETS AND BARS
Info. overload	Many of the variables that influence the processes are not measured in real time since they are obtained from laboratory.	**Huge amount** of process and material **data** from many individual unit operations.	Large data streams from processes, typically not directly associated with products, but with time. Many redundant sensor information.
Lack of meaning	Difficulty processing different data format and at different time scales	Complex relation among unit operations. Analysis of spectroscopy-based	Several variables that are not evaluated due to missing relation to processes. Principally inaccessible properties such as a continuous estimate for steel microstructure.
The need to act fast	Decisions taken are oriented to previous experiences of the operator. Lack of	Decisions between in-spec and out of spec (OOS) material have to be taken very fast. If only a small	Mechanical forming processes like rolling are quick, control systems

	digitization as sensors that measure bitumen content or viscosity and decision tools make decisions not optimal. Accordingly, digital systems retrofitting is the first step to be carried out in this type of plants.	amount of OOS material gets in the good product, the entire product has to be discharged. In advanced control, mitigation of OOS events is prioritized. Therefore, data from different units and sensors has to be collected, aligned and analysed in real-time.	require reaction times <100ms. Decision to estimate the outcome of a process, including potential OOS events, must be supplied sufficiently fast between two processes, <1min. Sensorial data must be assessed in (process) real-time.
What to remember	Much of the information is not collected or stored in computer systems. Information is not followed up. The decision depends on the expertise of the plant operator. There is no complete traceability of the production process. There is a gap between planning, production and logistics.	Recording of relevant data is crucial to fulfil regulatory constraints. The critical quality attributes of raw material, intermediates and final products must be tracked throughout the entire processing line. Process data must be linked to the drug product in order to guarantee that product quality is within the specified limits.	Product-oriented data must be collected, associating the gathered process data univocally to one product. Together with quality information/KPI per product, which then act as labels, a relation between quality and process streams can be established. Product interrelationships must be stored, because sometimes the quality of one product depends on the process data of products processed shortly before.

Table 1: Impact of cognitive biases in CAPRI use case sectors

Based on the work in D2.2, following figure illustrates the ToBe architecture for process monitoring and control realized in a cognitive plant.

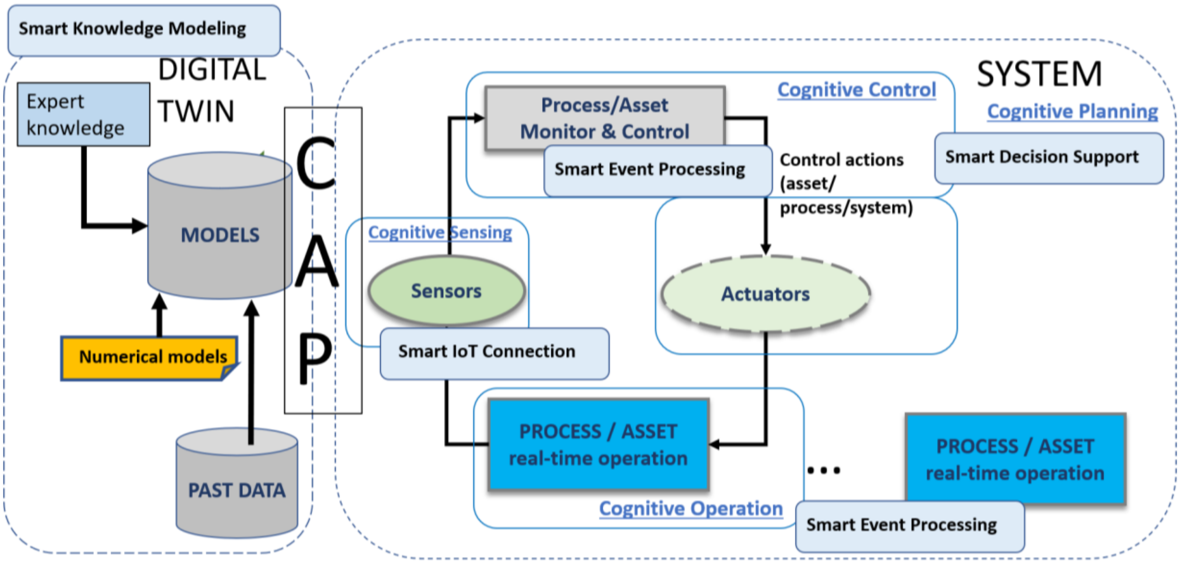


Figure 7 Cognition-driven process monitoring and control loop (cognitive plant)

In order to efficiently realize this system, there is a need for defining three meta-processing steps, as depicted in the following figure, reflecting the human cognition process, which was one of the main inspirations for the cognitive plant architecture.

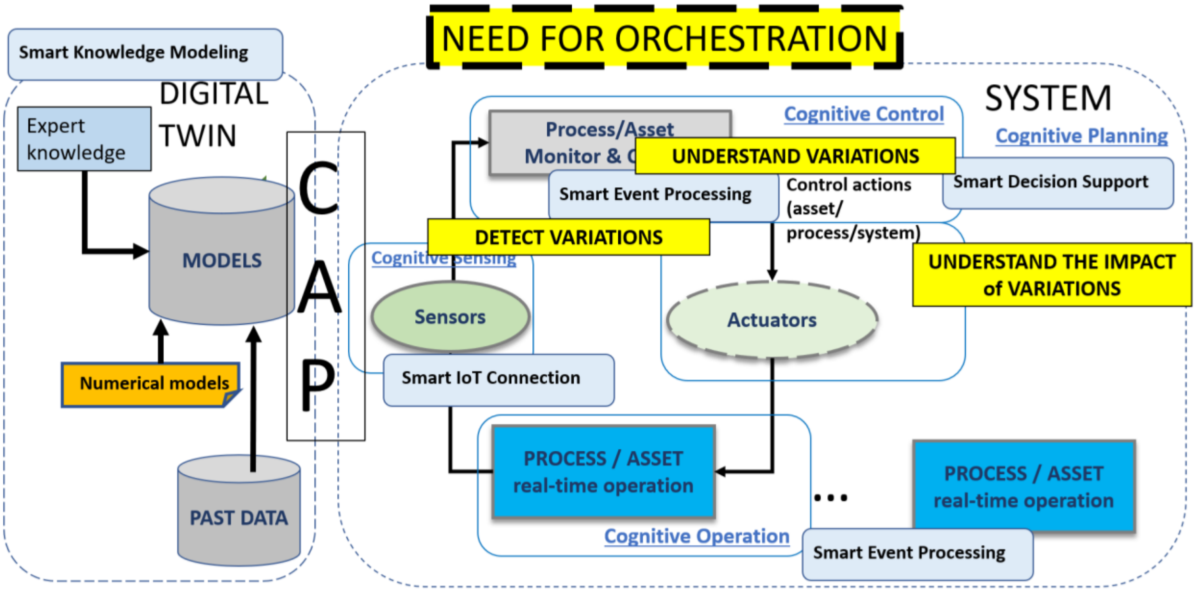


Figure 8 Cognition-driven process monitoring and control loop (cognitive plant)

Very briefly:

1) Detect variations is related to the ability to sense any variation in the high-dimensional real-time input streams
2) Understand variations is about interpreting the variations using available data and models and discovering their root causes (why the variations have appeared)
3) Understand the impact of variations is related to predicting the behaviour of the system if the variations will be not eliminated (what if the variations will be not stabilized).

As shown in figure, there is a need for an orchestration functionality, which will enable the synchronization of the meta-processing steps.

3.2.1 Role of Edge Computing

Across a range of industries, and specifically in the industrial automation vertical, there is broad agreement that the deployment of modern computing resources with cloud native models of software lifecycle management will become ever more pervasive. Placing virtualized computing resources nearer to where multiple streams of data are created is well established. It is the path to address system latency, privacy, cost and resiliency challenges that a pure cloud computing approach cannot address. This paradigm shift was first coined by Cisco in 2012, then incorporated in IIC Reference Architecture, under the label “fog computing” and progressively morphed into what is now known as “edge computing”.

The full potential of this transformation in both computing and data analytics is far from being realized. The mission critical requirements are much more stringent than what the cloud native paradigms can deliver. This is especially true because mission critical applications have five specific requirements:

- Heterogeneous hardware – Typical industrial automation settings have different architectures, as well a variety of compute configurations on the floor;
- Security – The security requirements and their mitigations vary from the device to device and need to be handled carefully;
- Innovation – While some of the industrial applications can continue with the legacy paradigm of remaining the same for over a decade, most of the industrial world now additionally requires modern data analytics and monitoring of applications in their installations;
- Data privacy – as in other areas of IT, data permission management is increasingly complex within connected machines and needs to be managed right from the origination of the data
- Real-time and determinism – the real-time determinism provided by controllers remains critical to the safety and security of the operation.

For these reasons current SotA is converging to “mission critical edge” solutions, that would incorporate requirements typical of embedded computing (security, real-time and safe, deterministic behaviors), into modern networked, virtualized, containerized lifecycle management and data and intelligence rich computing.

This kind of distributed system architecture is the enabler to securely consolidate, orchestrate and enrich with the fruits of data analytics and artificial intelligence (AI) the many poorly connected, fragmented and aging subsystems controlling today’s industrial environments. In this we can highlight the following points:

- Distributed and interconnected, mixed criticality capable, virtualized multi-core computing nodes (system of systems);
- Networking support that includes traditional IT communications (e.g., Ethernet, Wi-Fi) but also deterministic legacy field busses, moving towards IEEE time sensitive networking (TSN), and public and private 4G/5G, also moving towards determinism;
- Support for data distribution within and across nodes, based on standard middleware (OPC UA, MQTT, DDS, and more) will also strive towards determinism;
- Distributed nodes which have to be remotely managed and software will be delivered and orchestrated as virtual machines (VMs) and containers, the model of modern cloud native microservices.

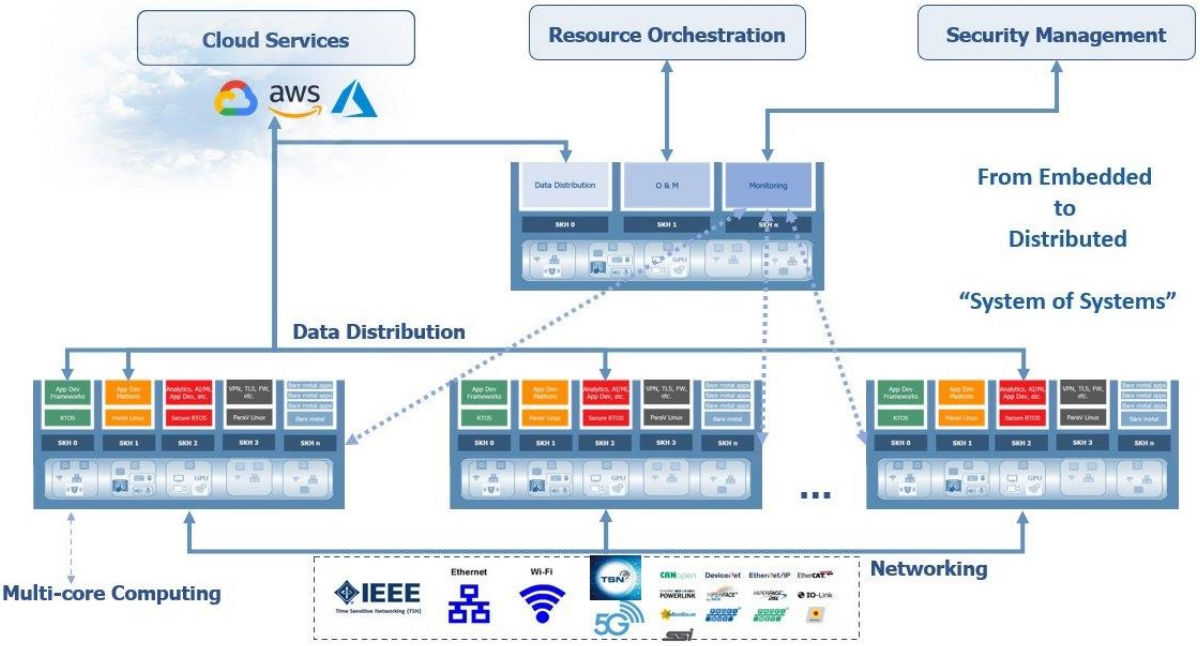


Figure 9 Embedded to Distributed "system of systems" for future industrial environments. (Source: Lynx Software Technologies)

The evolution of the architecture on the industrial automation floor plays a major role for the realization of the full mission critical edge paradigm, providing a unified and uniform infrastructure, going from the machine, through the industrial floor, and into the telco edge and cloud, enabling a fundamental decoupling between hardware and software. Applications, packaged as VM and, more and more, as containers, can be lifecycle-managed and orchestrated across all the layers of this infrastructure. Subsystems controlling today's physical environments can be effectively and securely consolidated, orchestrated and enriched with the fruits of data analytics and artificial intelligence.

The figure below shows how the infrastructure would look when the mission critical edge would be deployed, embedded into the operational technologies area of the factory. There are a distributed set of nodes, some very close to the plant, some far away. Effectively this is like a distributed datacenter, yet contains a far more heterogeneous, interconnected virtualized set of computing resource which can host the applications where needed and when needed. These will be deployed in the form of virtual machines and containers orchestrated from the cloud or locally.

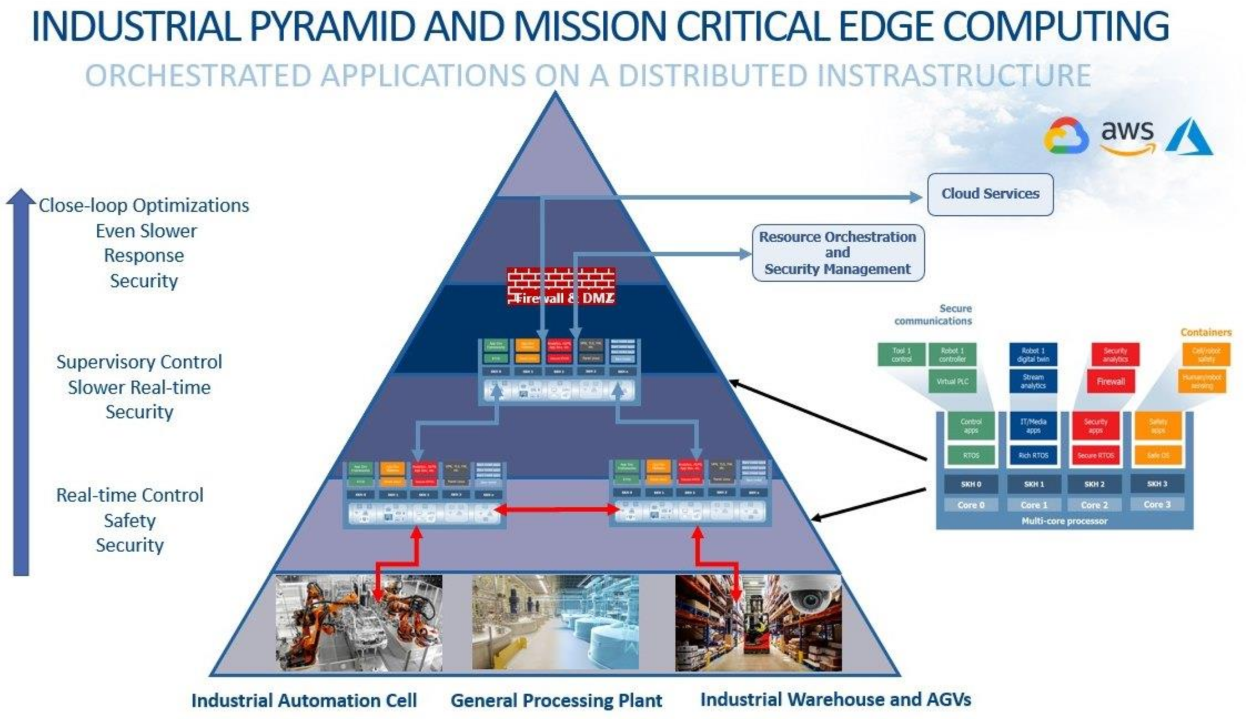


Figure 10 Foreseen distributed architecture on the industrial automation floor, with mission critical edge deployed. (Source: Lynx Software Technologies)

3.3 Cognitive Automation Platform (CAP)

The support for digital cognitive plants in process industries means transforming the traditional plant in a cognitive plant, integrating it with custom-made connectors for the data ingestion, both data in motion coming from the shopfloor and data at rest coming from the legacy systems, implementing smart algorithms capable to improve manufacturing processes, at sensor-control-operation-planning level. For this CAPRI will leverage the Smart Industry FIWARE Architecture, which is based on industry standards and Open Source components, supporting several communication protocols such as MQTT, CoAP, OPC UA, that facilitate the development of smart applications for all production processes.

CAPRI Reference Architecture will be based on the Industrial Internet Consortium RA, that enables the transformation of traditional process factories into cognitive process plants, integrating cognitive tools in order to cover all industrial automation levels such as control operation and planning.

The Cognitive Reference Architecture will provide the glue that integrates business intelligence, operational dashboards, complex event processing, machine learning algorithms, managing the security of the data and systems and guaranteeing the sovereignty.

The Cognitive Reference Architecture will be the reference point for the digital transformation of cognitive plants starting from the collection of data, moving through the reasoning of the data, up to the exploitation of data with configurable dashboard.

While FIWARE for Industry can provide reference implementations to most of the functional components foreseen in the IIRA specifications, some improvements need to be developed to provide cognitive functions at sensor, edge and cloud level. For this aim the FIWARE for Industry RA, will be enhanced in CAPRI by adding cognitive functions to compose the powered by FIWARE Cognitive Process Plants Blueprint, and making easier the development of cognitive components as well as the integration of them in a more comprehensive way.

The resulting CAPRI RA and Blueprint will enhance the capability to reach the expected benefits and impacts on the proposed use case, but also to apply many of the cognitive solutions to other sectors.

From the Asphalt domain, many solutions could be applied to process plants in cement and mineral sectors. Steel production domain has similarities with copper, aluminium, zinc and other metal-processing industries.

For both cases final modelling is different and must be tailored case-by-case, but the overall method, sensorial equipment, training and algorithms can be directly transferred, and this facilitate the wider impact of the CAPRI project, making possible the exploitation into different markets and sectors.

From the Pharma domain, the resulting solutions could be shifted to any other industry based on processing of powder raw material. The investigated unit operations feeding, granulation and fluid bed drying are widespread in industry. All sectors utilizing the mentioned process steps can benefit from implementation of the concept of in-line monitoring of attributes like concentration, blend uniformity and granule size distribution, the modelling of mentioned unit operations and the development of the cognitive control concept. Food industry is especially related to pharma processing and faces the same challenges. The concepts developed for the tableting process can be employed in future for many consumer products such as dishwashing tabs, etc.

Last, but not least, cognitive automation platform, will be virtually applicable to any industrial sector, although the focus of CAPRI is process industry sector with mixture of continuous and batch processing steps.

3.3.1 Cognitive Automation Platform Roadmap

WP2 is the starting point of an evolution from the Requirement gathering and the definition of a basic methodology and architecture for digital transformation of process industries into cognitive process industries.

After that the development and validation of cognitive technology solutions working independent starts in WP3 in industrial laboratory conditions (TRL5). Then WP4 deals with the development and demonstration (TRL6) of such cognitive technologies working integrated together within the cognitive automaton platform (CAP) in the real environments provided by the use cases.

Based on the generated CAP (WP4) and the cognitive tools (at Planning, Operation – WP4 - and Control – WP3), its integration and demonstration as prototypes (TRL7) at the three use cases (WP5) will show the potential for improved performance in cognitive production plants and will provide feedback of the lessons learnt for the final delivery of Digital Transformation Methodology (WP2).

During the whole project life-cycle, a strategy and deliveries for an effective dissemination of CAPRI innovations (WP6) will enable exploitation of project's results aiming at commercial results, and dissemination to next generation SPIRE employees (WP7), whether from training at premises, education integrated in existing curricula or lifelong learning programmes.

At the same time, for the project duration, CAPRI partners will cooperate with other projects with the aim to enhance cross pollination of results based on lessons learnt and guaranteeing that the data generated in the course of the project will be accessible and reusable.

During the project different versions of the software and hardware components will be developed:

a) The project will start with a comprehensive analysis of use cases for a functional description (requirements of individual solutions and architecture) ready at M8 (MS1),

b) Then the **alpha-version** at M18 (MS3) involving **cognitive automation architecture** and **validation of the technology** for cognitive individual solution modules (from planning to sensing) in laboratory industrial relevant environment,

c) after that, in M24, the **integration of technology** for cognitive solutions within the CAPRI architecture for planning and operation and its relation with preliminary interfaces with CAPRI automation platform,

d) following, the **beta-version** at M36 involving the demonstration of cognitive technologies at each functional level (planning, operation and control) integrated into CAPRI cognitive automation platform (CAP).

A period of 6-month validation of CAPRI prototypes at each of the use cases for demonstration and replication purposes is expected.

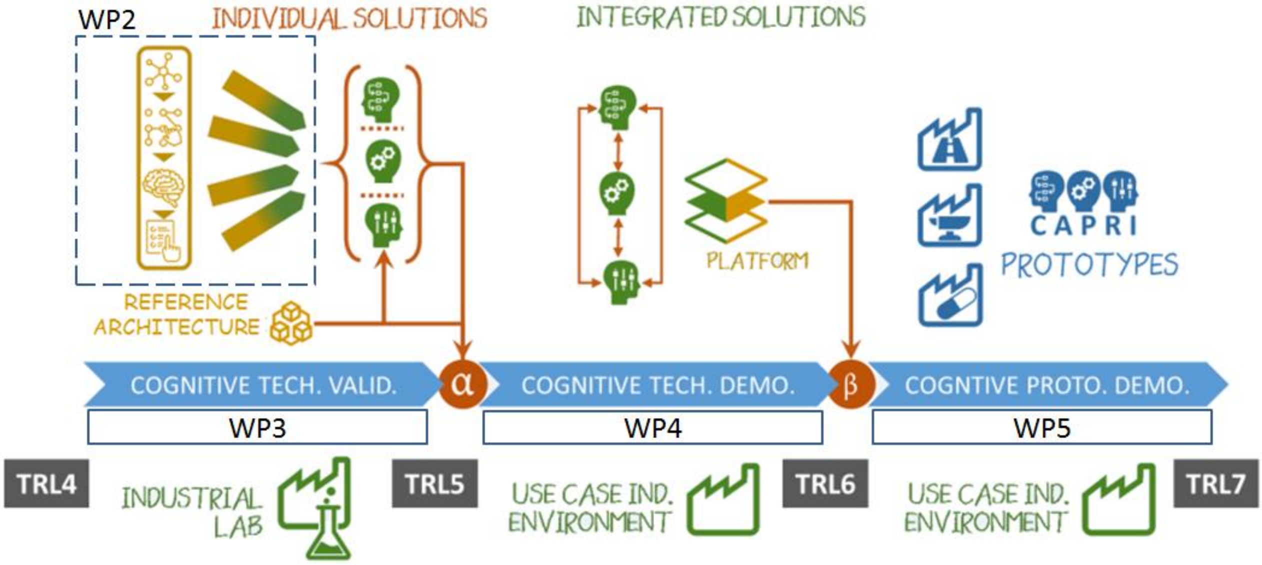


Figure 11 Planned Development roadmap

4 Relevant Reference Architectures

4.1 Industrial Internet Reference Architecture (IIRA)

The Industrial Internet Reference Architecture (IIRA) has been published by the Industrial Internet Consortium (IIC) in the document "The Industrial Internet of Things Volume G1: Reference Architecture" (The Industrial Internet of Things Volume G1: Reference Architecture Version 1.9, 2019) and contains architectural concepts, vocabulary, structures, patterns and a methodology for addressing design concerns. The document identifies the fundamental architecture constructs and specifies design issues, stakeholders, viewpoints, models and conditions of applicability defining a framework by adapting architectural approaches from the ISO/IEC/IEEE 42010-2011 Systems and software engineering—Architecture description standard.

This international standard outlines the requirements regarding a system, software, and enterprise level architecture. The ISO/IEC/IEEE 42010 standard recommends identifying the perspectives of the various different stakeholders that can be: system users, operators, owners, vendors, developers, and the technicians who maintain and service the systems. The aim is to describe system properties as seen from their viewpoint. Such properties include the intended use and suitability of the concept in terms of its implementation, the implementation process itself, potential risks, and the maintainability of the system over the entire lifecycle.

Essentially, the IIRA attempts to identify the most important and common architecture concerns. It then provides an architectural template and methodology that engineers can use to examine and resolve design issues. In addition, the template and methodology suggest ways of addressing the top concerns, allowing designers to glean insights by examining architecture patterns, helping Industrial Internet of Things (IIoT) system designers to avoid missing important architecture considerations and this also helps them to identify design gaps of missing important system functions or components.

IIoT system's architecture and representation are outlined in the IIRA in a generic way, keeping a very high level of abstraction, in order to be applicable to the depth and breadth of every small, medium and large enterprise of several sectors (Manufacturing, Transportation, Energy, Healthcare…). This is why it fits also with CAPRI needs of defining a common methodology that covers requirements from different companies: it is adopted as a fundamental support tool in CAPRI Reference Architecture definition, providing standard-based concepts and highlighting aspects to take care about.

The core of the IIRA's methodology lies in a set of system conceptualization tools called viewpoints that enable architects and engineers to identify and resolve key design issues. Thus, the IIRA design starts with defining the shapes and forms of an Industrial Internet of Things Architecture by starting with the viewpoints of the stakeholders. These IIRA's viewpoints are arranged in a particular order to reflect the pattern of interactions that occurs between them, because the decisions from a higher-level viewpoint impose requirements on the viewpoints below it. In this sense, the IIRA is a layer model that takes into consideration four different viewpoints: business, usage, functional, and implementation.

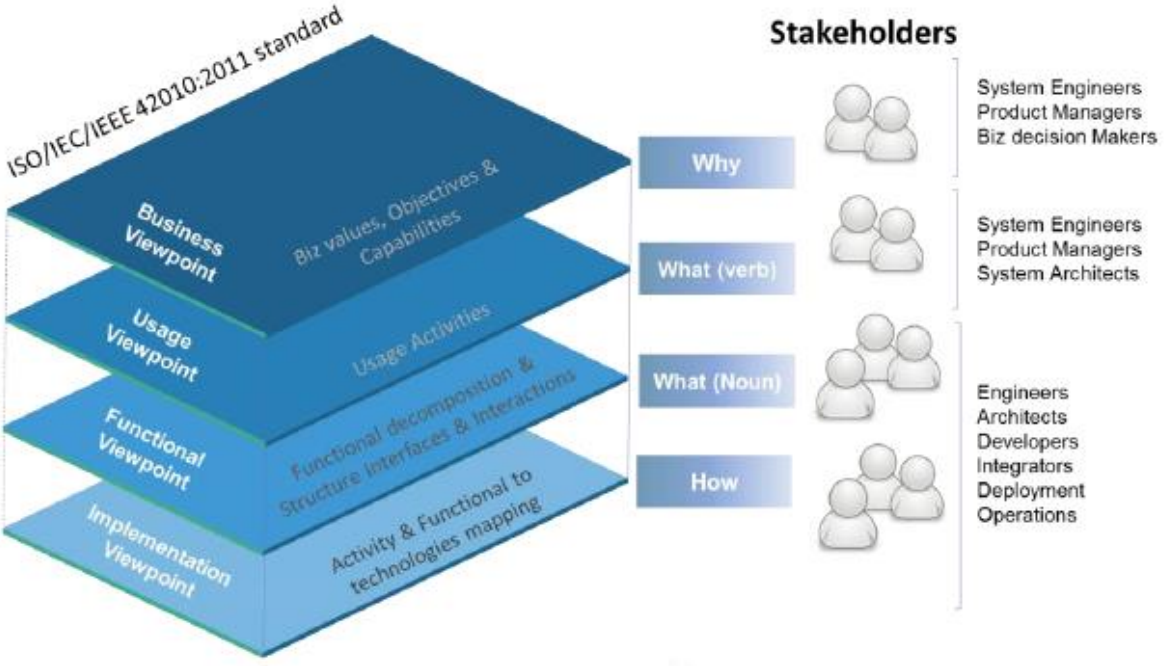
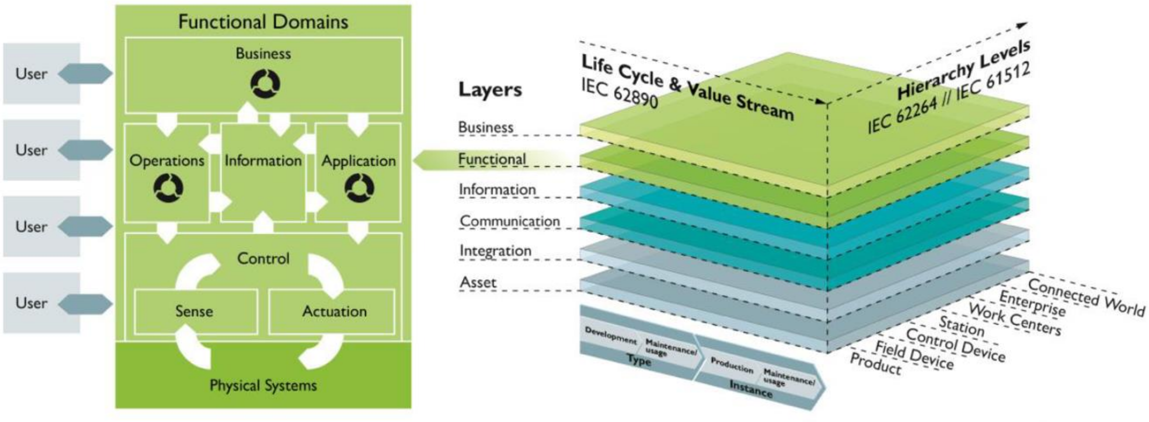


Figure 12 IIRA Architectural Framework

IIRA also focuses on the capabilities from the perspective of the software and their business processes. Each of the four viewpoints outlined in IIRA can be compared with the respective layers on the vertical axis of RAMI 4.0 (described in detailed in next paragraph); RAMI 4.0 supplements the model with the axes 'Lifecycle' (with types and instances) and 'Hierarchical Levels.'



Quelle: Plattform Industrie 4.0

Figure 13 The IIRA functional domains vs. RAMI 4.0 model

In addition to IIoT system architects, the plain language of IIRA and its emphasis on the value proposition and enablement of converging Operational Technology (OT) and Information Technology (IT) enables business decision-makers, plant managers, and IT managers to better understand how to drive IIoT system development from a business perspective.

4.2 Reference Architectural Model Industrie 4.0 (RAMI4.0)

The Reference Architectural Model Industrie 4.0 (RAMI 4.0) was developed by the Platform Industrie 4.0 in 2015 and focuses on the IoT and Cyber-Physical Systems (CPS) in the industrial manufacturing domain. RAMI4.0 is a three-dimensional model, which describes the Industrie 4.0 space and organises the lifecycle/value streams and the manufacturing hierarchy levels across the six layers of the IT representation of Industry 4.0.

One of the main objectives once adopted is to be able to communicate the scope and design of the system, to further collaboration and integration with other relevant initiatives by framing the developed concepts and technologies in a common model.

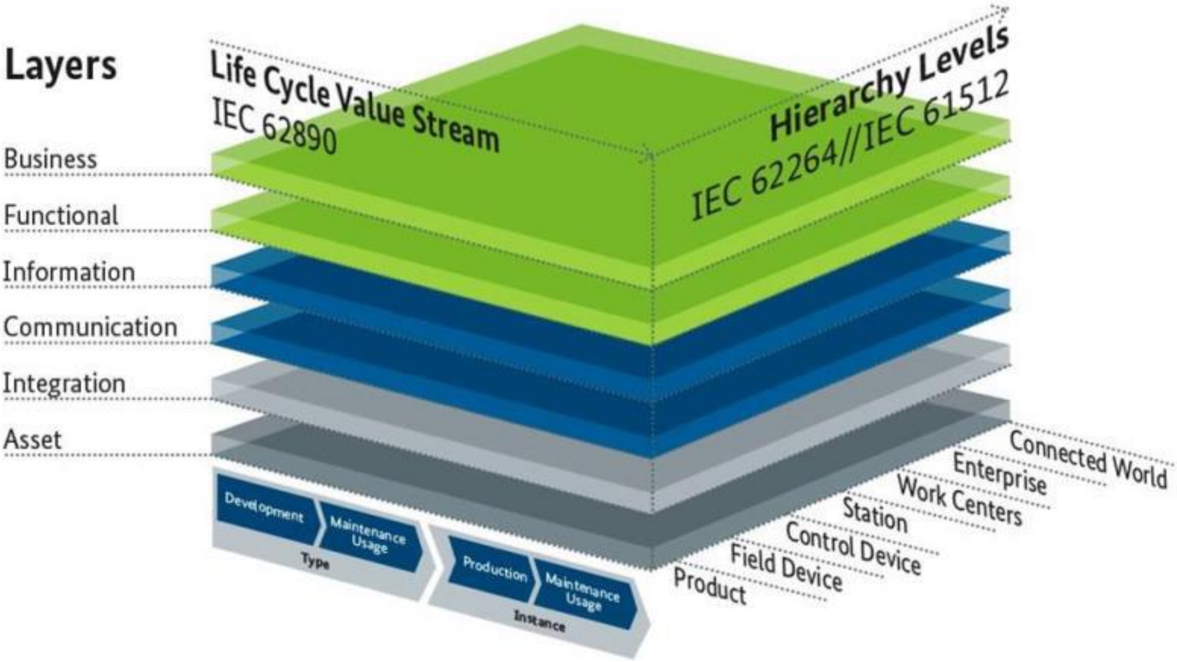


Figure 14 The three dimensions of the RAMI 4.0 (Source: Platform I4.0 and ZVEI)

The three-dimensional matrix can be used to position standards and describe use-cases. It addresses integration within and between factories, end-to-end engineering and human value-stream orchestration. This model is complemented by the Industrie 4.0 components and both have been described in DIN SPEC 91345. (Reference Architecture Model Industrie 4.0 (RAMI4.0) - DIN SPEC 91345:2016-04, 2016)

In RAMI4.0, each component consists of six layers. Starting with the lowest layer, the structure consists of asset, integration, communication, information, functional and business and represents a layered IT system structure, as shown in the figure below.

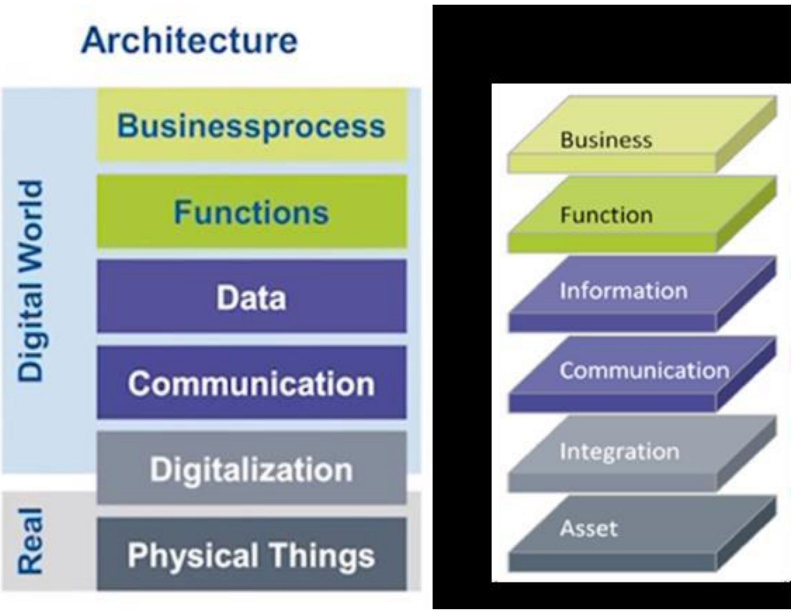


Figure 15 The IT Layers of RAMI 4.0 (Source: Platform I4.0 and ZVEI)

The function of each layer is:

- The asset layer describes physical components of a system, for example production equipment, product part, sensors, documents, as well as humans. For every asset represented in this layer there must be a virtual representation in the above layers. Among the physical assets, this layer includes the digital interface with humans and the relationship to elements in the integration layer.
- The integration layer deals with easy to process information content and can be considered as a bridge between the real and the IT world. It contains all elements associated with the IT, including field buses, HMIs, necessary to implement a function, as well as the properties and process related functions required to use an asset in the intended way and generates events based on the acquired information.
- The communication layer is responsible for the standardized communication between integration and information layer. Therefore, it performs transmission of data and files and standardizes the communication from the Integration Layer, providing uniform data formats, protocols and interfaces in the direction of the Information Layer. It also provides services to control the integration layer.
- The information layer holds the necessary data in a structured and integrated form and provides the interfaces to access this structured data from the functional layer. It is responsible for processing, integrating and persisting the data and events, as well as for describe the data related to the technical functionality of an asset. It can be considered the run-time environment for Complex Event Processing (CEP) where rule-based (pre-) processing of events takes place, data APIs and data persistence mechanisms. So, events are received from the communication layer, transformed and forwarded accordingly.
- The functional layer describes the logical and technical functions of an asset providing a digital description of its functions and a platform for horizontal integration of various functions; it also describes the business model mapping, business processes which can be adjusted based on inputs from the functional layer, providing models with runtime data of processes, functions and applications.
- The business layer is in charge to orchestrate the services provided by the functional layer. It maps the services to the business (domain) models and the business process models. It also models the business rules, legal and regulatory constraints of the system. The processes to ensure of the economy are located on this level.

In order to represent the Industry 4.0 environment, the functionalities of IEC62264 have been expanded to include two new levels, at the bottom, the “product” (both the type and the instance, through the entire lifecycle) which are active elements within the production system due to their ability to communicate. They provide information on their individual properties and necessary production steps. This new level perfectly fits with CAPRI project purposes. Some of the Cognitive Solutions that will be developed will take advantages of the newest technologies to gather data and information from the product, enhancing productivity and efficiency of the entire process. It allows thus, a continuous connection between production field and the higher levels of the architecture, establishing the “product” as a functional level. At the top there is the “connected world”, which represents its outer networks or the ecosystem, e.g.,, collaboration with business partners and customers, suppliers or service providers, as well as Internet-based services. This allows moving from the typical pyramid, with rigid hierarchical structures, to a composite of networked objects and systems as reflected in the figure below.

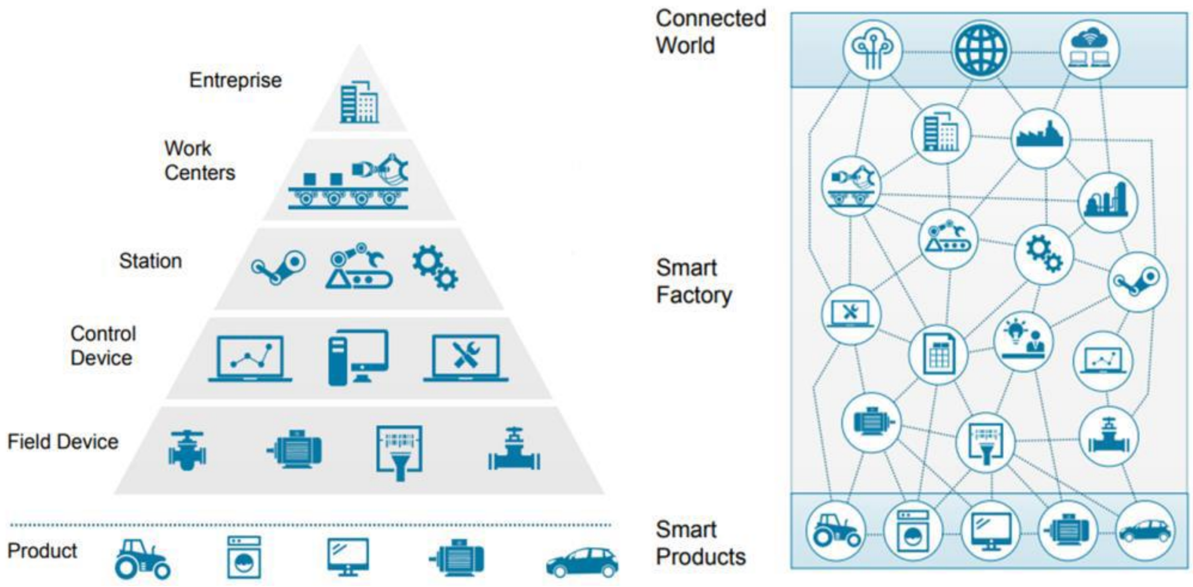


Figure 16 Hierarchy Levels of Industry 3.0 and RAMI 4.0 (Source: Platform I4.0 and ZVEI)

4.2.1.1 Asset Administration Shell (AAS)

The Asset Administration Shell (AAS) provides the basis for the development and the use of common standards in Industry4.0. The need of components, applications and devices to be able to communicate across domains, companies or countries is translated into the need of high interoperability of the Industry4.0 solutions. The main purpose of the AAS is to support and to help this transition defining a common way through which I4.0 solutions are linked together.

A key element of the AAS is the asset. An asset, in Industry4.0, it is more than a simple physical item. it represents in a holistic way all the information, sensors, software, resources and also human resources that belong to a company. An asset is then an identifiable entity with proper characteristics and status. AAS allows the assets in Industry4.0 to communicate with other assets of the same company or with others outside the company boundaries, in a I4.0 compliant manner.

The RAMI4.0 architecture and AAS are thus strictly interconnected, since one of the three dimensions of RAMI4.0 architecture entirely concerns the Life Cycle of an asset; on the other hand, AAS defines the way in which assets can be integrated into I4.0 structures, in accordance to their requirements. Inside the RAMI4.0, the standard IEC 63278 proposes interoperability and communication solutions between assets and the key capabilities can be summarised as follow:

- To structure, store and grant the availability of all the data, static and dynamic of an asset.
- To provide a standardised communication interface.
- To publish the asset's services to the outside, thus allowing its control via AAS.

4.3 FIWARE for INDUSTRY

FIWARE[6] aims to create a platform that uses a combination of technologies such as IoT, Big Data or Cloud architectures. FIWARE is open source and eases the creation of new applications in multiple verticals since it contains a rich set of components that can be deployed and connected in

---

[6] www.fiware.org

This project receives funding in the European Commission's Horizon 2020 Research Programme under Grant Agreement Number 870062

a compliant way.

FIWARE tries to distribute the data and vertical silos in many IoT based systems by using a horizontal layer so that it manages large-scale context information. FIWARE NGSI defines an information model and a RESTful interface can be constructed by context data providers and consumers. The latter can be realized by interacting with the Context Broker, which is a central component of FIWARE architecture. The Context Broker aims to enable the system to make updates and access to the current state of context.

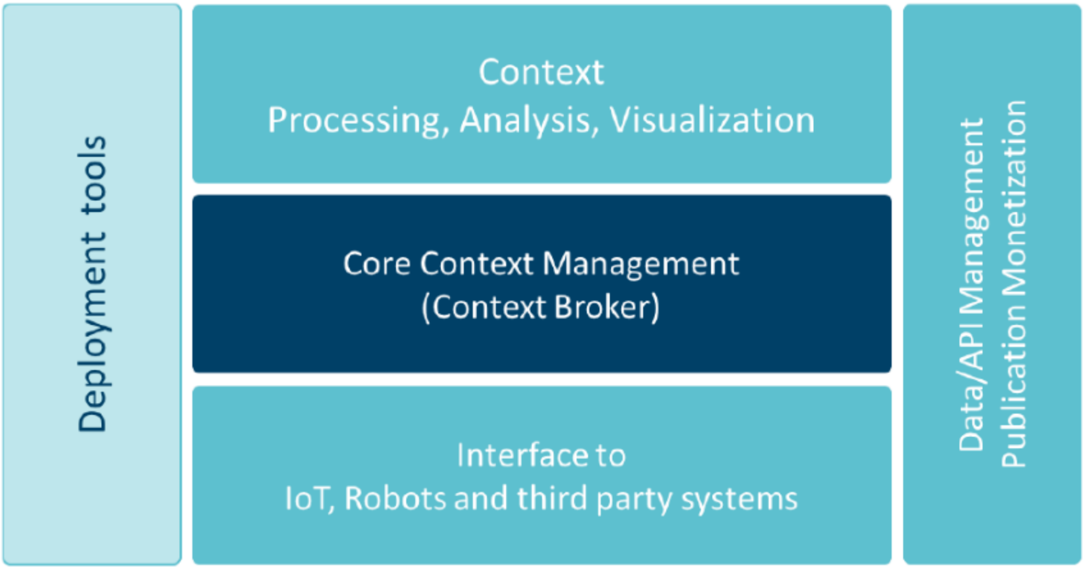


Figure 17 FIWARE overall architecture

FIWARE includes specifications of the interfaces of multiple Generic Enablers (GE) as depicted in the figure above. The FIWARE Catalogue is a curated framework of Open Source platform components which can be assembled together and with other 3rd Party platform components to accelerate the development of Smart Solutions.

The main and only mandatory component of any “Powered by FIWARE” platform or solution is the FIWARE Orion Context Broker Generic Enabler, which brings a cornerstone function in any smart solution: the need to manage context information, enabling to perform updates and bring access to context.

Building around the FIWARE Context Broker, a rich suite of complementary FIWARE components are available, dealing with:

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

The catalogue contains a rich library of components with reference implementations that allow developers to put into effect functionalities such as the connection to the Internet of Things or Big

Data analysis, making programming much easier. FIWARE core platform model facilitating IaaS and SaaS required of application domains, on this basis GEs applications achieve already defined standards, provide APIs for interoperability, represent application domains or design granularities. We can think to a GE as Macroscopes were highest level interface is a simple controller providing a wide and in scope view of operations (attributes control functions of a system); GEs from different domains are Macroscopes on the domain: implement abstract Macroscopes concretely and provide API access via REST HTTP to trigger GE behaviour. Modelling a GE is identified within UML use cases. GE specification have some properties:

- Addressing: IP address and port numbers
- Recognition: control syntax, parser, interpreter and semantic rules
- Multimodal: APIs, protocols, drivers
- Structured Data: XML, JSON, IC
- Formal Operation: state machine, dispatcher, DOM nodes
- Ad hoc Network Communication: HTTP/s, request methods, asynchronous, client/server, URIs
- Modular Design: object-oriented architecture, methods and functions, listeners, callbacks
- Behavioural: multithreaded, parallel, imperative, result combining, verifying, transformation, bidirectional coms
- Security: channel encryption, message encryption, authentication, authorization
- HCI: GUI, hardware interaction, multimodal UI, accessibility, human actors (Doctors, Patients, Staff, …)
- Interoperability: networked API server, configuration parameters, legacy system integrators, RPC, REST

The ETSI Industry Specification Group for cross-cutting Context Information Management (ISG CIM) (Industry Specification Group (ISG) cross cutting Context Information Management (CIM)) has just released a preliminary specification of an API considered to be cornerstone in the development of Smart Cities, Smart Agrifood or Smart Industry applications.

In any smart solution there is a need to gather and manage context information, processing that information and informing external actors, enabling them to actuate and therefore alter or enrich the current context. Group Specification CIM 004, referred to as the ETSI NGSI-LD API specification, defines a simple way to update or query context data within a Smart Application, including factors such as the source, meaning, licensing, or related information describing that data.

The ETSI ISG CIM has decided to give the name “NGSI-LD” to the Context Information Management API to reinforce the fact that it leverages on the former OMA NGSI 9 and 10 and FIWARE NGSI specifications, incorporating the latest advances from Linked Data. The FIWARE Context Broker component (Orion), the core component of any “Powered by FIWARE” platform or solution, provides the Open Source reference implementation of the FIWARE NGSI API and will evolve to work as open source reference implementation of the new ETSI NGSI-LD API specifications.

The implementations of NGSI-LD which are available are: Orion-LD[7], Scorpio[8], and Djane[9].

More specifically, the Orion Context Broker collects data from sensors, drones, vertical smart solutions and information systems. In this way, the broker breaks information silos. Sensors are connected to IDAS IoT Agents, so that they can handle many IoT protocols such as MQTT,

---

[7] https://github.com/Fiware/context.Orion-LD
[8] https://github.com/ScorpioBroker/ScorpioBroker
[9] https://github.com/sensinov/djane/

CoAP/OMA-LWM2M, OneM2M. Also, alternative IoT platforms can be used for this situation. Fast RTPS is used to interface ROS-2 robots, which is the main communication middleware in ROS-2. Different processing engines, such as Flink, Hadoop and Spark, are used in order to process historical data, so as to extract valuable insights or derive smart actions. Artificial Intelligence or Complex Event Processing functions can be used above the integrated processing engines. Wirecloud web mashup framework is used for Operating dashboards. Extended CKAN portal can offer to 3rd parties part of the current and historic context data. The API/Data access control functions enable access to the context data to parties that own certain privileges. The API management and business support layer can offer auditing of the system and monetize data access.

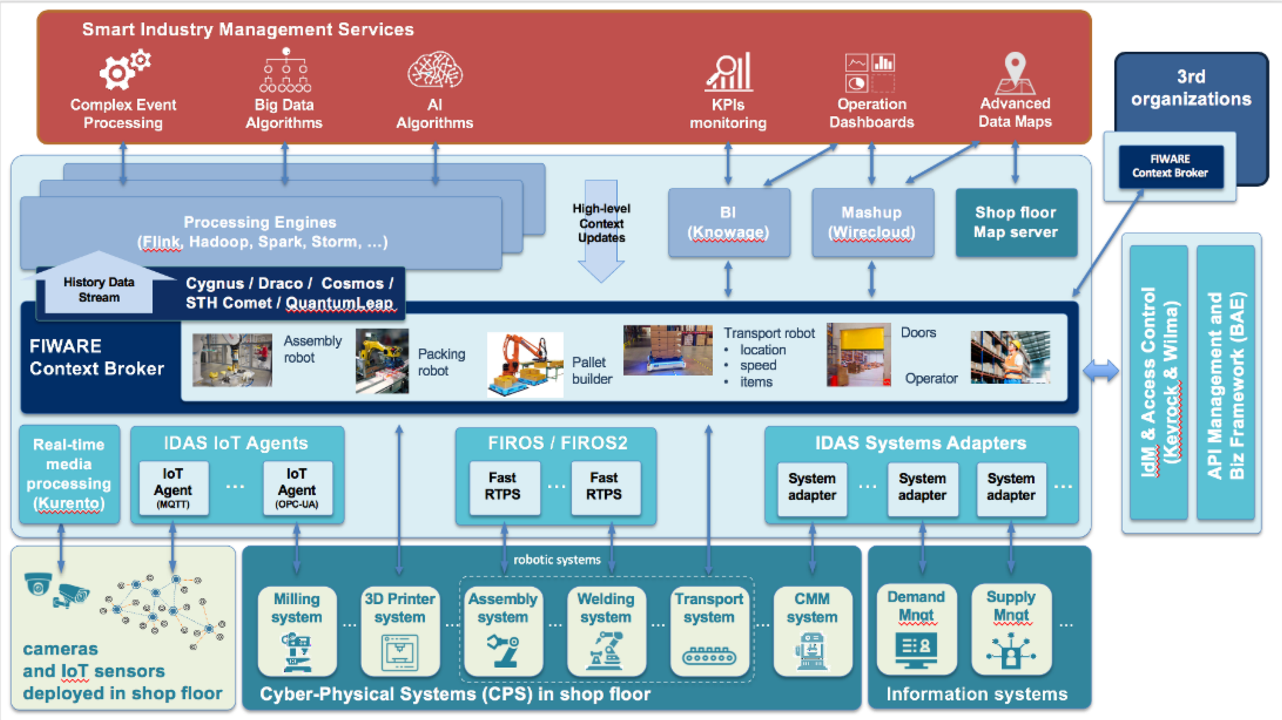


Figure 18: FIWARE Smart Industry Reference Architecture (v2020/middle)

The above picture the current snapshot of the continuously evolving FIWARE Smart Industry Reference Architecture, updated on a regular base by the FIWARE Smart Industry Mission Support Committee (chaired by ENGINEERING and NEC). Thanks to this strong collaboration with the FIWARE Industrial Community, CAPRI will be able to remain up to date on the novelties coming from the Community, as well as it will be able to influence this collaborative endeavour being among the first projects applying FIWARE technologies in the process industry.

4.4 Big Data Value Association (BDVA)

The BDVA Reference Architecture[10] is a reference framework made by the European BDVA (Big Data Value Association) that describes logical components of a generic big data system.

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 Public Private Partnership (PPP) on Big Data Value. In the context of this deliverable the most important information is related to the Reference Model, for the overall positioning of concepts and the technical priorities, for the relation with manufacturing scenarios. The BDVA Big Data Value Reference Model can be seen in the following figure.

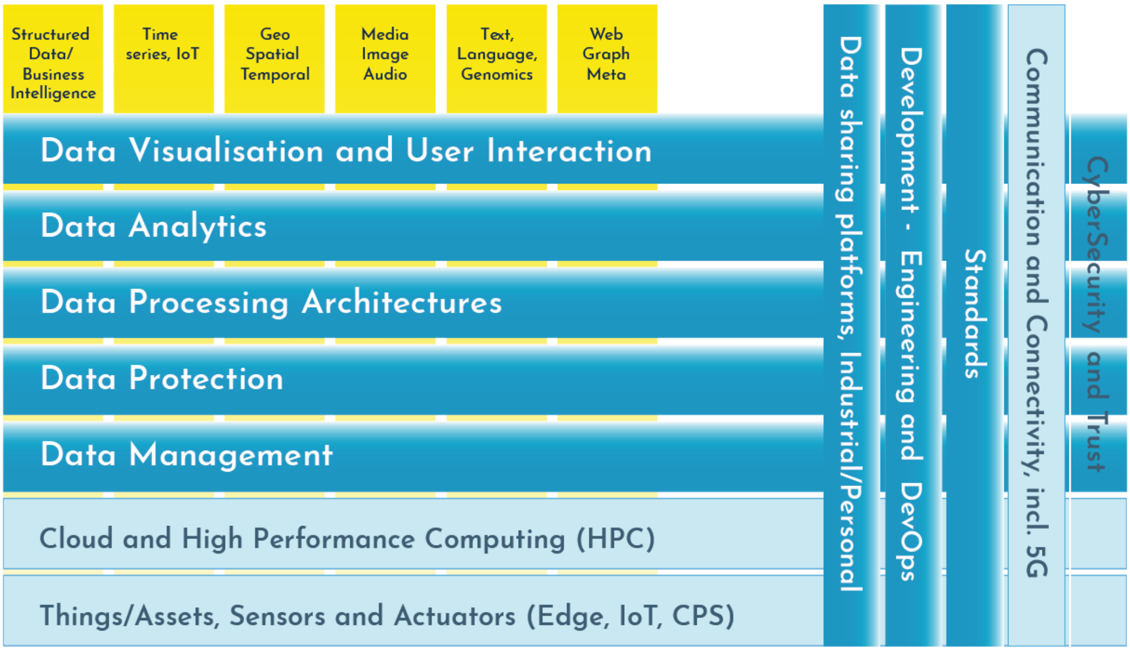


Figure 19 The BDV Reference Model

In fact, BDVA has proposed their initiative regarding a European Data-Driven Artificial Intelligence and their vision regarding AI and Big Data and how it can drive the European technology and economy[11]. To realize this vision, it will be necessary to address a number of challenges:

a) Data-driven AI-based solutions for the industry will require new business models.
b) Trust in AI and its results must be established; for example, one should be able to explain how AI applications came to a specific result ("Explainable AI"), which would foster responsible technological development (e.g., avoid bias) and enhance transparency in how and why an AI takes a decision.
c) It is necessary to develop an AI and Big Data ecosystem, by developing data for open AI platforms and overcoming the lack of data interoperability.
d) Fuse and develop a number of technologies, as a successful industrial AI relies on the combination of a wide range of technologies, such as advanced data analytics, distributed AI, and hardware optimized for AI.

---

[10] European BDVA Strategic Research and Innovation Agenda v4.0. (2017, October). Retrieved from http://www.bdva.eu/sites/default/files/BDVA_SRIA_v4_Ed1.1.pdf

[11] Data-driven Artificial Intelligence For European Economic Competitiveness and Societal Progress. BDVA Position Statement. (2018, November). Retrieved from http://www.bdva.eu/sites/default/files/AI-Position-Statement-BDVA-Final-12112018.pdf

Those technical priorities that are relevant as expressed in the BDV Reference Model are the following:

1. **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.
2. **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.
3. **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.
4. **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.
5. **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.

To this end, recently Big Data Value Association (BDVA) and euRobotics have launched the creation of the Data, AI and Robotics Association (DAIRO) PPP (public-private partnership)[12]. The wider purpose of objectives and activities are Big Data Value, AI, Data and Robotics update in objectives and activities.

This new association have objectives in order 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.

4.5 International Data Spaces Association (IDSA)

The International Data Spaces Association (IDSA) is the evolution of IDS (Industrial Data Space) which itself was an initiative lead by Fraunhofer ISST, in cooperation with ATOS, T-Systems, and the idea is promoted by the German Federal Ministry of Education and Research. IDSA is characterized by the focus on information ownership, with the aim of enabling clear and fair exchanges between data providers and consumers. To this end it suggests a reference distributed architecture that accomplishes this goal (IDS Reference Architecture Model Version 3.0.

Broadening the perspective from an individual use case scenario to interoperability and a platform landscape view, the IDS Reference Architecture Model positions itself as an architecture that links different cloud platforms through policies and mechanisms for secure data exchange and trusted data sharing (through the principle of data sovereignty). Over the IDS Connector, industrial data

---

[12] Joint Vision Paper for an AI Public Private Partnership (AI PPP). Brussels: BDVA –euRobotics. (2019). Retrieved from http://www.bdva.eu/sites/default/files/VISION%20AI-PPP%20euRobotics-BDVA-Final.pdf

clouds, individual enterprise clouds, on-premise applications and individual, connected devices can be connected to the International Data Space ecosystem (see Figure below).

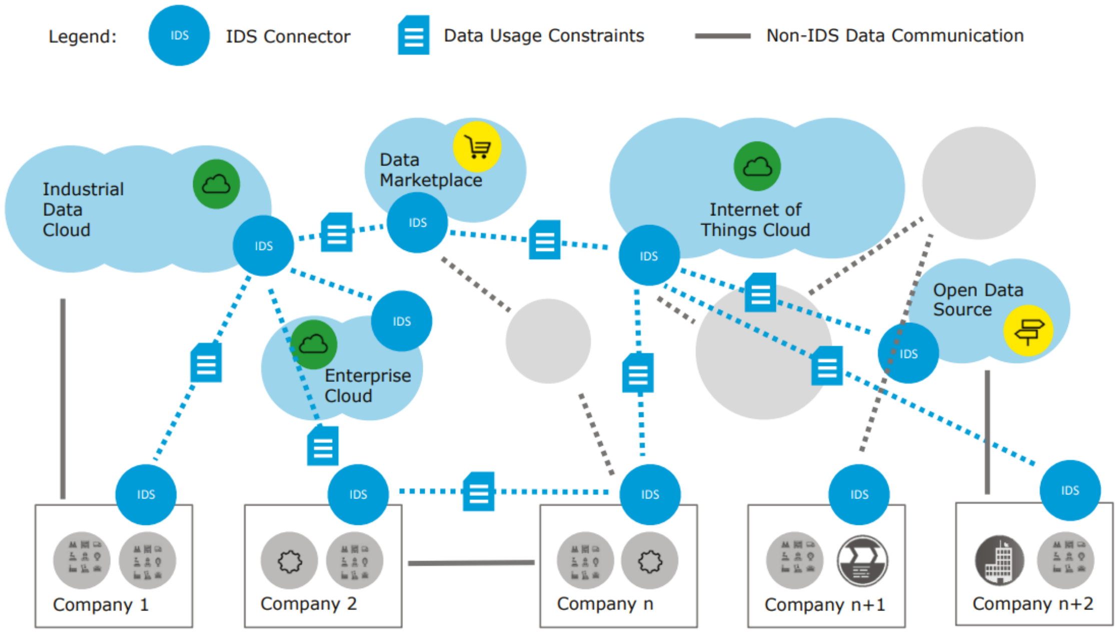


Figure 20 International Data Spaces connecting different platforms

This IDS Reference Architecture Model (IDS-RAM) is described using multiple layers, such as business, functional, process, information and system; between and common to all these layers are transversal functionalities that foster security, certification and governance, as illustrated in Figure 20. The Business Layer specifies and categorizes the different roles which the participants of IDS can assume, and it specifies the main activities and interactions connected with each of these roles. The Functional Layer defines the functional requirements of IDS, plus the concrete features to be derived from these. The Process Layer specifies the interactions taking place between the different components of IDS; using the BPMN notation, it provides a dynamic view of the Reference Architecture Model. The Information Layer defines a conceptual model which makes use of linked-data principles for describing both the static and the dynamic aspects of IDS' constituents. The System Layer is concerned with the decomposition of the logical software components, considering aspects such as integration, configuration, deployment, and extensibility of these components.

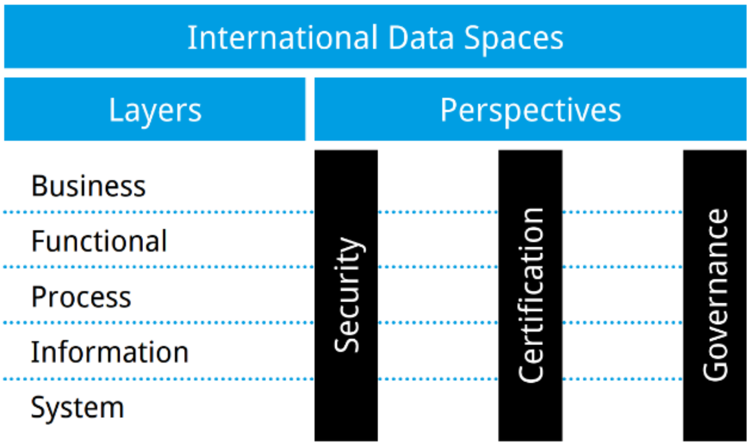


Figure 21 General structure of Reference Architecture Model

Comparing IDS to IoT-A ARM, the former focuses its specification of the roles for actors within the business layer that would govern the data flows between different domains or data spaces. As such, key participants (actors in the system) would be the Data Owner, Data Provider, Data Consumer, Data User or Broker Service provider. The complete landscape of roles, their functionalities and relationships result in a model depicted in the following Figure.

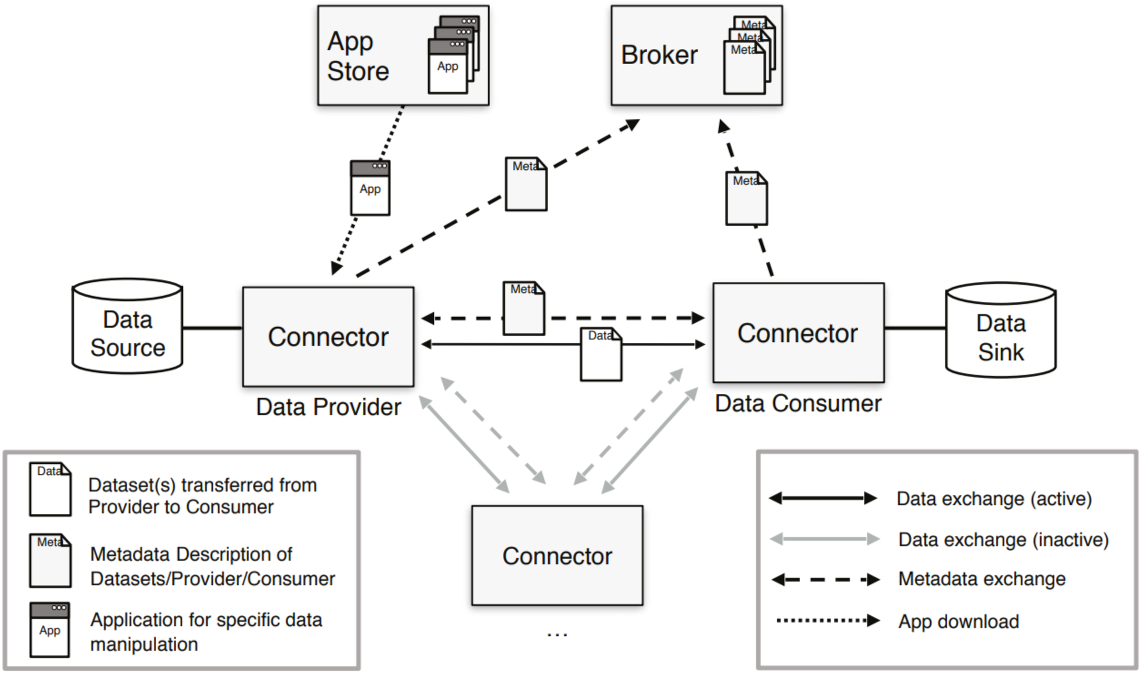


Figure 22 Interaction between technical components of IDS Reference Architecture Model

The **Connector** is the central technological building block of IDS. It is a dedicated software component allowing Participants to exchange, share and process digital content. At the same time, the Connector ensures that the data sovereignty of the Data Owner is always guaranteed. The **Broker Service Provider** is an intermediary that stores and manages information about the data sources available in IDS. The activities of the Broker Service Provider mainly focus on receiving and providing metadata that allow provider and consumer connectors to exchange data. The **App**

Provider role is optional in IDS, and its main role is to develop applications that can be used by both data providers and consumers in the data space. Applications are typically downloaded from the remote app store, and run inside the containerized connector.

Establishing **trust for data sharing and data exchange** is a fundamental requirement in IDS. The IDS-RAM defines two basic types of trust: 1) Static Trust, based on the certification of participants and core technical components, and 2) Dynamic Trust, based on active monitoring of participants and core technical components. For data sharing and data exchange in the IDS, some preliminary actions and interactions are required. These are necessary for every participant, and involve a Certification Body, Evaluation Facilities, and the Dynamic Attribute Provisioning Service (DAPS). Figure 22 illustrates the roles and interactions required for issuing a digital identity in IDS, and these interactions are briefly listed here:

1. **Certification request**: This is a direct interaction between a participant and an evaluation facility to trigger an evaluation process based on IDS certification criteria.
2. **Notification of successful certification**: The Certification Body notifies the Certification Authority of the successful certification of the participant and the core component. Validity of both certifications must be provided.
3. **Generating the IDS-ID**: The Certification Authority generates a unique ID for the pair (participant and component) and issues a digital certificate (X.509).
4. **Provisioning of X.509 Certificate**: The Certification Authority sends a digital certificate (X.509) to the participant in a secure and trustworthy way and notifies the DAPS.
5. **Register**: After the digital certificate (X.509) is deployed inside the component, the component registers at the DAPS.
6. **DTM Interaction**: The Dynamic Trust Monitoring (DTM) implements a monitoring function for every IDS Component, and DTM and DAPS then exchange information on the behaviour of the component, e.g.,, about security issues (vulnerabilities) or attempted attacks.

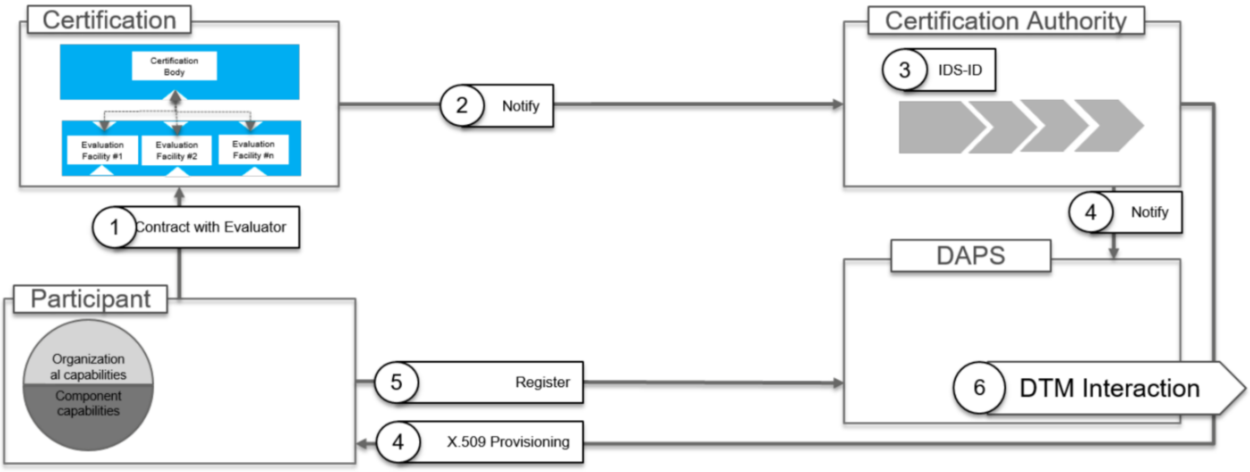


Figure 23 Interactions required for issuing a digital identity in the IDS

The IDS Reference Architecture contains an internal structure that is strongly supported by the containerization for the development of IDS connectors. It relies on IDS Communication Protocol to enforce security in data exchanges, as it is depicted in the figure below.

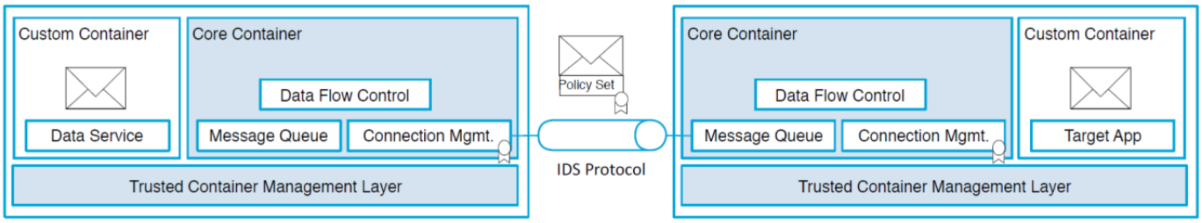


Figure 24 Enforcing security in data exchanges: the IDS Communication Protocol

To sum up, the security implications that guarantee reliable and trusted transfer of information between independent entities in IDS are the following:

- **Secure communication.** The concept of Trusted Connector is introduced as depicted in the figures above.
- **Identity Management** for identification/authentication/authorization enhancing. There is use of certificates issued by a Certificate Authority (CA).
- **Trust Management** that uses Cryptographic methods such as PKI (Public Key Infrastructures).
- **Trusted Platform** for trustworthy data exchange, which defines the minimal requirement for Security Profiles that should be verified by IDS connectors. It also defines the capacity to perform integrity verification of the rest of the involved connectors.
- **Data Access control**. IDS defines authorization criteria based on the previously defined Security Profiles.
- **Data Usage Control**. IDS checks and regulates that data processing is according the intended purposes defined by the original data owner.

4.6 Other relevant projects in SPIRE

Under the same SPIRE umbrella of CAPRI project, where it is needed to develop new technologies to realise cognitive production plants, with improved efficiency and sustainability, by use of smart and networked sensor technologies, intelligent handling and online evaluation of various forms of data streams, several other SPIRE projects can be found where several Reference Architectures will be needed to be applied.

4.6.1.1 COGNIPLANT (Cognitive platform to enhance 360° performance and sustainability of the European process industry)

COGNIPLANT project will develop and demonstrate an innovative approach for the advanced digitization and intelligent management of the process industries. This approach will be based on a novel vision to data monitoring and analysis, that will make the most of the latest developments on advanced analytics and cognitive reasoning, coupled with a disruptive use of the Digital Twin concept to improve Production plants' operation performance by up to 68% in real time control of the productive environment, 65% in quality control of the final products and 70% in response time to uncontrolled incidents.

The concept will be implemented by four end-users from four different SPIRE industries, one chemical industry in Austria, one aluminium refinery in Ireland, one concrete manufacturing industry in Italy and one metal industry in Spain. The COGNIPLANT solution will provide a hierarchical monitoring and supervisory control that will give a comprehensive vision of the plants' production performance as well as the energy and resource consumption. Advanced data analytics will be applied to extract valuable information from the data collected about the processes and their effect on the production plant's overall performance enabling to design and simulate operation plans in digital twin models based on the conclusions. As a result, optimal operation plans will be obtained that will improve the performance of those cognitive production plants. In addition, the project will demonstrate the positive impact derived from the implementation of COGNIPLANT solution that will allow industries reducing their CO2 emissions up to 20%.

A training strategy will be designed to provide a comprehensive framework for the dissemination of the project outcomes and a clear understanding of the new solution for the employees of the SPIRE sectors.

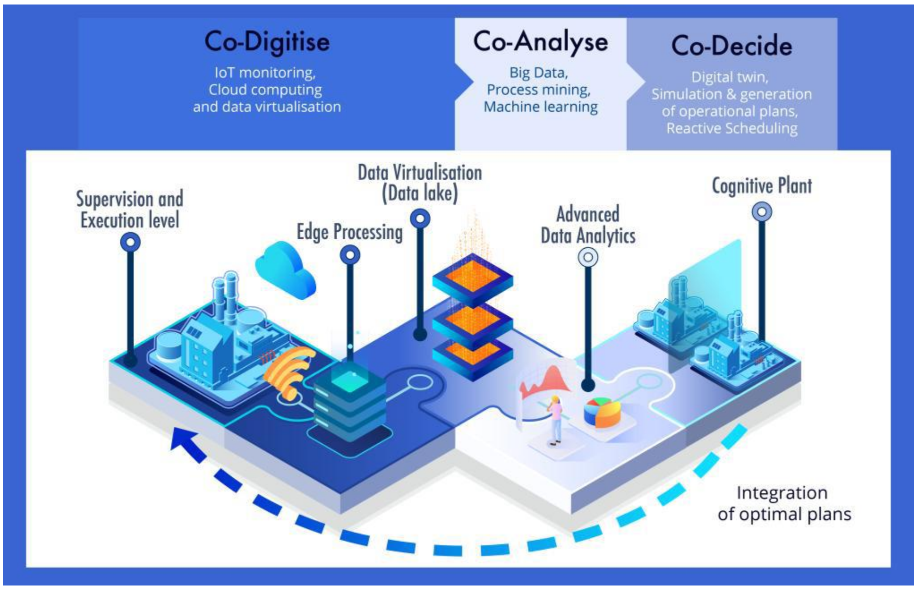


Figure 25 Cogniplant framework

Therefore, COGNIPLANT project will be based on three main layers:

- Sensing and Data Virtualisation level ► "Co-Digitise": collect and structure the data from the different sensors and equipment for its further analysis.
- Advanced data analytics level ► "Co-Analyse": data processing, application of advanced methods of process mining, big data, data mining, etc.
- Virtual Model and simulation level ► "Co-Decide": digital twin, decision making, generation of operational plans, prescriptive general and edge processing.

4.6.1.2 COGNITWIN (Cognitive plants through proactive self-learning hybrid digital twins)

COGNITWIN aims to add the cognitive element to the existing process control systems and thus enabling their capability to self-organise and offer solutions to unpredicted behaviours.

To achieve the objectives of the project, there are six industries and seven research groups from seven European nations, each of whom will bring their expertise in data analytics and pattern recognition which are going to be at the heart of the COGNITWIN solution platform. The set-up of the platform includes a sensor network that will continuously monitor and collect data from various plant processes and assets which will be stored at a database. This data will be used to develop a digital twin of the process and will also be used to develop models with cognitive capability for self-learning and predictive maintenance which will lead towards optimal plant operations.

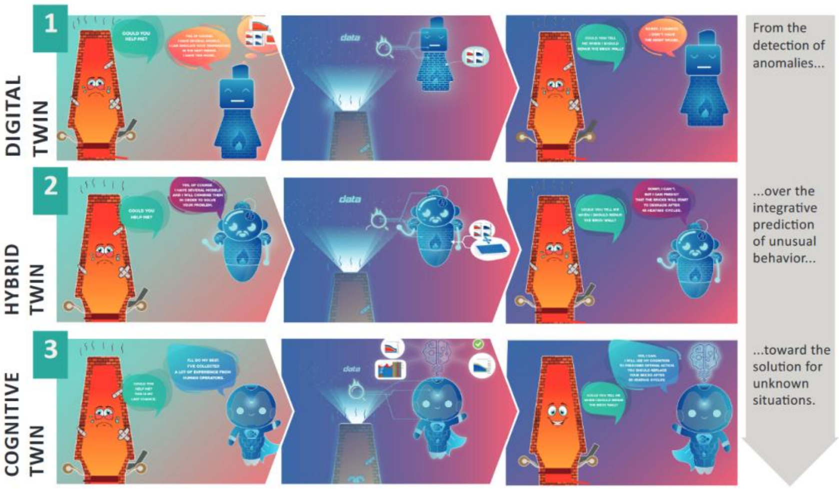


Figure 26 Cognitwin layers

Specific COGNITWIN objectives can be grouped into the following issues:

- COGNITWIN for Industry Process Excellence: Show improved performance in cognitive production plants by a technology demonstration of fully digitalized pilots.
- Cognitive Digital Twins for Cognitive Retrofitting: Enabling an efficient and well-defined approach for “cognitive augmentation” of physical assets, processes and systems for Cognitive Digital Transformation in Process industry.
- Hybrid Twins for Optimised Process Performance by hybrid models that combines first principle and data-driven models and use machine learning, AI and the connected data bases to pro-active forecast and communication, as well as self-learning by recognition of patterns in the data.
- COGNITWIN Interoperability Toolbox as a Service: A reference architecture for the cognitive elements including of Big Data, Databases, IoT, Smart Sensors, Machine Learning, and AI technologies that realizes hybrid modelling, self-adaptivity and cognitive recognition, leveraging/extending the existing work into relevant communities.
- COGNITWIN for increasing European Technology Dominance: Ensure the dominance of the Europe in technologies related to cognitive plants, thereby influencing the further development of Big Data, Databases, IoT, Smart Sensors, Hybrid Modelling, Machine Learning and AI technologies in relevant communities, focusing on the capabilities of the developed technologies for creating new generations of self-adaptive and cognitive algorithms and models.
- COGNITWIN for SPIRE: Ensure the knowledge transfer of results and experiences from the COGNITWIN project to the SPIRE Process Industry community, focusing on active participation in the new SPIRE DG7 Digitalisation group and in SPIRE organized events.

4.6.1.3 FACTLOG (Energy-aware Factory Analytics for Process Industries)

Particularly in process industries that cognition can improve the behaviour of a complex process system. One of the main expectations of the use of digital twins is to give the capability to observe and monitor the behaviour of their respective physical twins. In order to make it happen, it is needed to combine digital twins, which are driven by domain models (i.e., knowledge), with the models derived from data (i.e., experience). In order to realize it, a real-time processing layer is needed where observations (i.e., events), knowledge and experience interoperate to understand and control the behaviour of a complex system (i.e., cognition). FACTLOG offers such a layer and aims at deploying and adjusting it to several process industries.

By incorporating different pipelines of machine learning and analytical tools at different levels (from machines to process steps and from processes to the whole production plant), FACTLOG enables the realization of the Cognitive Factory as an ensemble of independent but intertwined ECTs, that are (i) able to self-learn, and thus to effectively detect and react to anomalies and disruptions but also to opportunities that may arise, (ii) enjoy a local or global view of operations and (iii) are capable for short-, mid- and long-term reasoning and optimization.

FACTLOG is driven by several specific, yet indicative, business cases in the process industry and focuses its innovation regarding analytics, AI and optimization on the deployment and assessment of coherent Enhanced Cognitive Twins for the specific sectors represented in FACTLOG and even for the plants in which it will be pilot tested and evaluated.

It will be implemented by a just right consortium of twenty partners, five of which are manufacturers and a further three represent manufacturing clusters. Technology and scientific contributors from leading academic institute and focused ICT vendors (mostly SMEs) bring in all necessary knowledge and innovation.

Figure 27 Factlog manifesto

FACTLOG specific objectives include:

- To define the Cognitive Factory framework model.
- To enhance factory automation.
- To deploy model-driven analytics both at the lower level.
- To deploy data-driven services.
- To develop robust (re-)optimization methods and tools.
- To offer its fully-fledged toolset ‘as-a-service’ in an easy to adopt and deploy manner.
- To establish proof-of-concept.
- To enhance collaboration with academia and ongoing research projects within and beyond the SPIRE PPP.
- To offer measurable improvement in plant performance specifically for process industries.

4.6.1.4 HyperCOG (Hyperconnected Architecture for High Cognitive Production Plants)

HyperCOG addresses the full digital transformation of the process industry and cognitive process production plants through an innovative Industrial Cyber-Physical System (ICPS). It is based on commercially available advanced technologies that will enable the development of a hyperconnected network of digital nodes. The nodes can catch outstanding streams of data in real-time, which together with the high computing capabilities available nowadays, provide sensing, knowledge and cognitive reasoning to the industrial business.

Furthermore, HyperCOG is deeply grounded in the last advances in Artificial Intelligence such as modelling for twin factories, decision-support systems for human-machine interaction and augmented reality for industrial processes visualization. It pursues self-learning from the process in order to deal with the typical dynamic fluctuations of the industrial processes and global optimization.

The objective is to increase the production performance while reducing the environmental impact by reducing the energy consumption and the CO2 emissions thereof. Society will get profit of this project not only throughout the environmental impact, but through the lifelong learning of workers and vocational training for digitisation, and the available training modules for youth at schools such as ESTIA technological institute or U-PEC University.

The breaking-edge system proposed in HyperCOG project will be validated on the productivity and environmental impacts, replicability and usability aspects on three use cases belonging to the SPIRE scope such us SIDENOR (steel making), CIMSA (cement), and SOLVAY (chemical) use cases.

Figure 28 HyperCOG application fields

HyperCOG specific objectives aim at the following issues:

- Development of a platform that converts manufacturing industries into more flexible environments, capable of continuously adapting themselves to changing conditions. Current industrial systems are normally implemented following a hierarchical structure making them rigid and less efficient at making good and fast decisions. HyperCOG will build on existing robust core functionalities with modern digital technologies to improve this decision-making process.
- Implementation of advanced data analytics for extracting knowledge from production databases to optimise operations and support engineering management activities. This will be achieved using analytical concepts, dedicated software and IT systems which evaluate the data of the company, production data and the market development. HyperCOG aims to detect and isolate behaviour patterns inherent in the data which are related to production irregularities in order to suggest connections with specific causes and propose possible corrective or preventive actions.
- Development of a decision support system to make the best possible decision in a specific situation. Operators will receive information about errors in manufacturing and be provided with steps to solve the error and recover the operation.
- Leverage cybersecurity concerns about cyber-physical systems and Internet of Things devices as a business enabler. HyperCOG will deliver cybersecurity as a layer of protection, improving business efficiency and productivity through security and privacy by design, raising organisational resilience through situational knowledge by threat intelligence based preventive methods and leveraging on data analytics.
- Development of strategies for training and re-skilling human resources. The project will identify and implement professional competencies triggered by new scientific know-how and technological changes, and enable these competencies in an industrial context to address new professional needs.

4.6.1.5 INEVITABLE (Optimization and performance improving in metal industry by digital technologies)

The INEVITABLE Innovative Action aims to realize a fully digitalized monitoring technology for an optimized and improved performance of manufacturing processes. Use cases from the energy and resource intensive sectors steel and nonferrous metals are addressed, whereas the considered manufacturing sites are in Slovenia, Austria and Spain covering primary and secondary steelmaking and investment casting of nonferrous metal alloys.

The focus of INEVITABLE is to develop high-level supervisory control systems for different production plants and on the demonstration in operational environment (TRL 7) to enable autonomous operation of the processes based on embedded cognitive reasoning. Key Performance Indicators will be defined related to resource consumption and product qualities. Based on that, a full digital transformation of the plants will be done including acquisition, storage, processing and analytics of data streams, furthermore, communication and automation, and finally, standardization of relevant data interfaces. Predictive models will be developed being combined with smart and networked sensor technologies to correlate process parameters with quality indicators of the manufactured products. The models will be tested in offline mode on the one hand, and in online mode on the other hand by means of comprehensive plant trials at the industrial partner sites. Dissemination activities will transfer the knowledge throughout the SPIRE sectors. The industrial partners are supported by scientific partners with excellent competences in the field of digitalization.

INEVITABLE will improve the capabilities for reliable and real-time control logics of final product properties and process efficiency to increase the flexibility of plant operators. Improved and flexible production performance is expected with a simultaneous reduction of resource consumption and CO2 emissions contributing to a more competitive and sustainable metallurgic industry within the EU.

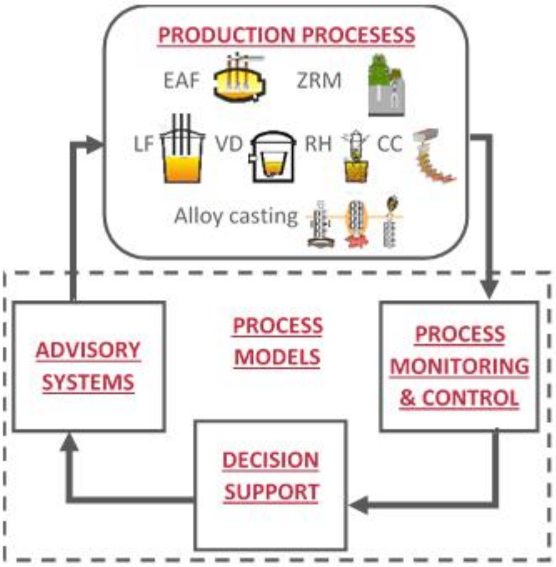


Figure 29 INEVITABLE process

The main ambition is to exceed the level and functionality of traditional process automation and control systems by applying the functionalities of Digital Factories and Industry 4.0 concepts.

The INEVITABLE project partners will develop and upgrade various technologies:

- Electric Arc Furnace (EAF) models and optimization framework,
- System for supervision, optimisation, diagnosis and condition monitoring of Cold Rolling mills (ZRM),

- Online advisory system for process control and optimization in clean steel intelligent manufacturing,
- Predictive process models for refining and continuous casting of liquid steel,
- Smart cognitive applications to support decisions in controlling the nonferrous casting process.

4.7 Other relevant projects applying AI in the industrial domain

In this section we describe some most relevant projects which apply AI in the industry.

4.7.1 AI REGIO

The AI REGIO project aims at filling the three major gaps currently preventing AI-driven DIHs from implementing fully effective digital transformation pathways for their Manufacturing SMEs: at policy level the Regional vs. EU gap; at technological level the Digital Manufacturing vs. Innovation Collaboration Platform gap; at business applicative level the Innovative AI vs Industry 4.0 gap.

- VANGUARD LEARN step. I4MS has developed in the recent years a methodology for DIH life cycle management, including a classification and critical analysis of their service portfolio (also adopted in the DIH.NET[13] CSA) and methods for measuring digital maturity and identifying opportunities for new services and cross-border collaborations. However, at the moment, just in very few cases such assets have impacted the regional strategies and policies as, for instance, implemented by VI regions. VANGUARD on the other hand is very well connected with regional and local Industry 4.0 initiatives especially for training, awareness creation and engagement of SMEs. However, at the moment such initiatives are quite independent each other and interoperability and collaboration prevented by deep differences in terms of conceptual frameworks, standards and KPIs measurement systems. AI REGIO will align VANGUARD digital transformation policies and programs with I4MS methods and tools, in order to create an interoperable network of AI DIHs, which will go beyond the 30 cases selected in the AI DIH Network[14] initiative.
- VANGUARD CONNECT step. Within I4MS, the MIDIH project has developed in the recent years an Innovation and Collaboration platform (DIHIWARE[15]) which is currently under experimentation on more than 20 DIHs. However, its penetration in local/regional Industry 4.0 initiatives is still in its infancy with the exception of few cases in Italy (Lombardy and Piedmont). VANGUARD on the other hand includes in its constituency strong regional networks which since years drive and materialize the EU innovation strategies. The 4 Motors for Europe[16] is one example, where the strongest four EU regions regularly meet and take important shared decisions about common policies including digital transformation ones. However, at the moment no digital platform is able to support such collaboration and innovation processes, including stakeholders communities, digital for a and innovation platforms. AI REGIO focusses on the 4 Motors for Europe movement as the most advanced movement for cross-regional collaboration and extends it to the whole VI regions and beyond, thanks to the deployment of the I4MS Community and Innovation / Collaboration platform
- VANGUARD DEMONSTRATE step. I4MS has recently involved hundreds of SMEs from all Europe in its “open calls” initiatives for innovative application experiments in several domains of robotics, HPC modelling and simulation, CPS/IOT, additive manufacturing. Very advanced and complete open platforms (and marketplaces) have been developed and potentially are ready to the go-to-market. However, no specific initiative is currently in place for Innovative AI experiments

---

[13] https://dihnet.eu/

[14] https://www.ai-dih-network.eu/

[15] As presented in the 6th EU DIH meeting on April 3rd https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=58544

[16] http://www.4motors.eu/en/

and Phase III IAs just marginally addressed the AI Industry 5.0[17] challenge especially regarding human acceptance. VANGUARD on the other hand includes pilot plants and often lighthouse plants strongly linked with national/regional initiatives and born from infrastructural funding opportunities which are not accessible by I4MS. This is also an assurance of multiplication and exploitation of results in less advantaged regions. Regarding AI, several EU regions are now involved in the AI Coordinated Plan and its monitoring action called AI Watch[18]. However, one the one side no decisions have been so far taken about Open Digital Manufacturing Platforms for AI and the presence and involvement of SMEs in these initiatives seems quite limited at the moment. AI REGIO aims at materializing common EU policies on AI and their monitoring actions in concrete reference architectures and Open Platforms for AI at the same time preserving EU values such as expressed by GDPR and Data Sovereignty principles and involving SMEs thanks to an intense program of open calls.

- VANGUARD COMMERCIALISE step. I4MS IAs have recently devoted more attention to these phases by collecting and publishing advanced catalogues and marketplaces of solutions especially in the domain of IoT and Modelling / Simulation. Of course, AI solutions marketplaces are not ready yet, but apart from that, the success of such technology platforms outside of H2020 is quite limited, especially because of lack of customers and buyers' communities. VANGUARD on the other hand can mobilize huge communities of potential customers and users of such solutions, but often low visibility is given to products and services originated from outside the specific regional borders. *AI REGIO aims at exploiting existing marketplaces of solutions and platforms by linking them with regional smart specializations and local policies (metadata and linked data) towards an interoperable common EU Data Space for innovative AI solutions for manufacturing).*

One of the important terms is "Collaborative Intelligence" envisaging two bidirectional collaborative interaction channels "Humans Assisting Machines" (train machines to perform certain tasks; explain the outcomes of those tasks and sustain the responsible use of machines) and "Machines Assisting Humans" (amplify our cognitive strengths and physical capabilities; interact with customers and employees; embody human skills to extend our physical capabilities). As discussed during the recent workshop organized by CIAO at EBDV Forum in Vienna and entitled "*Data-driven Collaborative Intelligence in Smart Manufacturing and Robotics*"[19], by taking the family environment paradigm, the former interaction is similar to a Parent-Child family relationship, where parents aim at transferring as much of their knowledge to children (train), at understanding their viewpoint and establishing a positive constant dialogue with them (explain) and at taking a collaborative and not punitive attitude when they make mistakes (sustain). *Humans therefore should see Machines as their own creatures and establish a personal affective relation with them: the more machines are perceived by humans as their unperishable digital heirs, the more AI will be accepted in next generation workplaces*. Caregiver-Elderly relationship is instead the family functional model when cognitive capabilities are insufficient (amplify), extensive and intensive multimodal interactions are needed (interact) and physical capabilities are inadequate (embody). *Humans therefore should see Machines as their capability multipliers, making it possible unprecedented and innovative activities at cognitive, physical and relational level*. According to the same HBR report, the following five principles are implementing Collaborative Intelligence in Industry, each focusing on one specific factor: i) reimagine business processes with an Ethics-by-Design perspective; ii) embrace experimentation/employee involvement following Inclusivity and non-Discriminatory principles; iii) actively direct AI strategy towards safety, Comfort and Wellbeing of employees; iv) responsibly collect data by adopting Privacy-by-Design principles; v) redesign work to incorporate AI and cultivate AI-related employee Skills and Professions.

---

[17] See the new DG RTD Unit F.5 https://ec.europa.eu/info/sites/info/files/organisation_charts/rtd_organigram_01_10_2019_en_hq.pdf

[18] https://ec.europa.eu/knowledge4policy/ai-watch_en

[19] https://www.european-big-data-value-forum.eu/program/data-driven-collaborative-intelligence-in-smart-manufacturing-and-robotics/

Main commonality is related to the MIDIH Platform for developing data-driven industrial solution. It is based on the Open Source frameworks, Apache and FIWARE, as presented in the following figure.

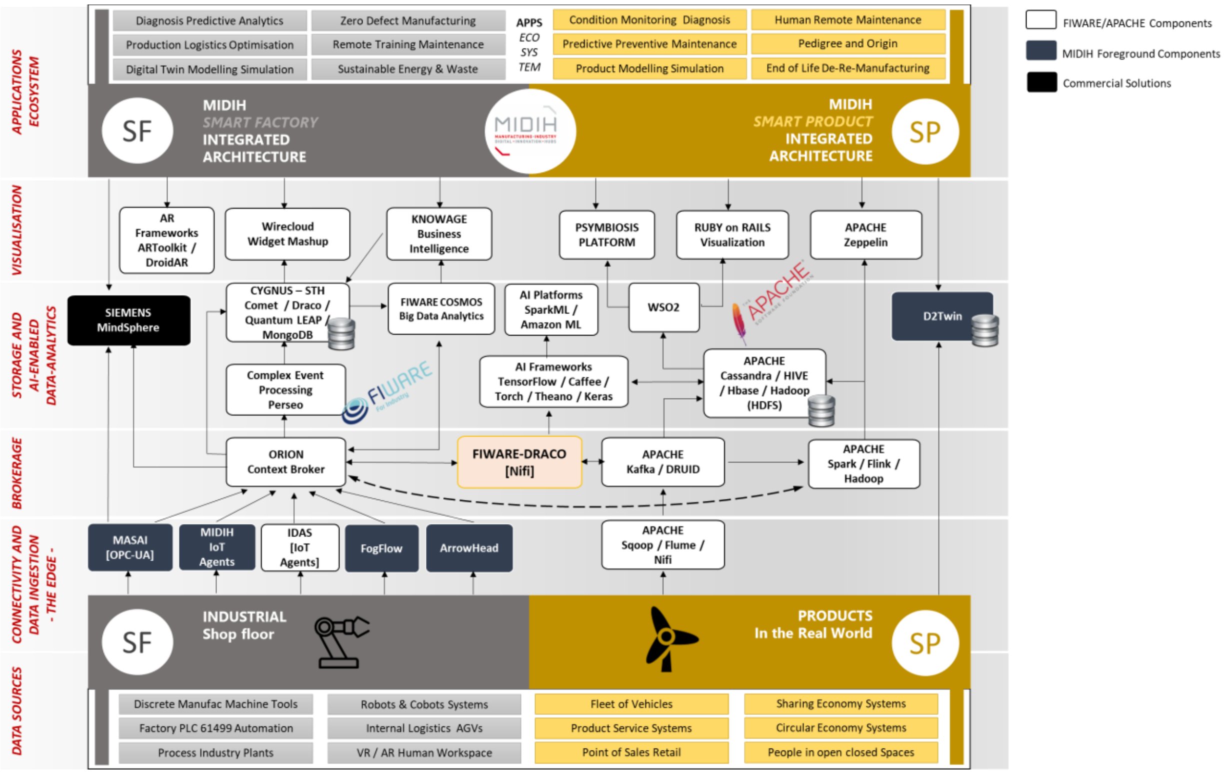


Figure 30 MIDIH Reference Architecture

This architecture is taken as the basis for the development of the new reference architectures, focusing on specific aspects of the industrial use cases to be supported with corresponding platforms.

4.7.2 AI4EU

AI4EU project[20] ("A European Artificial Intelligence on demand platform and Ecosystem") aims to make available to its users resources based on AI technology that facilitate scientific research and innovation, analyze future research needs in AI and create an observatory of ethics that ensures an AI focused on human being. One of the main objective is to build the first European Artificial Intelligence On-Demand Platform and Ecosystems.

The Vision for AI4EU on-demand platform is to build a platform which allows the AI community to publish and exchange AI assets and skills. By connecting innovation hubs (or DIH) with developers, the AI4EU on-demand platform will serve as a channel that gives access to AI resources to all related communities. AI4EU Platform will represent a focal node to access to AI resources, components, dataset… stored on the platform or distributed by other platforms, to shape news application, prototypes and components, to develop interaction and matchmaking with people, corporation, organization that can share your main interest and to collectively upgrade the platform content and governance. The AI4EU on-demand platform will be an enabling tool that empowers European innovators to speed up in developing AI-based innovation. By designing the AI4EU on-demand platform in a way that it is of mutual interest for all European communities, a flourishing interaction place connecting innovators (demand) with technology providers and researchers (supply) will be established. This will create the basis to align and build upon current AI research and innovation instead of continuous repetition and double work ignoring valuable AI assets already developed.

---

[20] https://www.ai4eu.eu/

The AI4EU platform will be the operational glue for many AI-related initiatives that have defined their vision / strategic documents by supporting the access and exchange of AI assets needed for boosting AI-based innovation. For instance, a company developing an application based on face recognition can rely on the AI4EU platform to access a competitive technology to build their solution instead of developing their own framework.

Current plans for AI4EU platform are to have second version of the platform by 2020 (when AI4PI would kick-off) and another improved version by 2021(at the middle of the project) together with a sustainability plan to go beyond the actual project itself. CAPRI will collaborate with AI4EU platform along three main avenues:

i) Use trustworthy AI resources (algorithms, software frameworks, development tools, components, modules, data, computing resources, etc.), explainable and verifiable from the beginning of the project;
ii) Provide AI resources from CAPRI use cases to AI European stakeholders through AI4EU repository[21];
iii) Being part of the platform community, exchanging knowledge with stakeholders, creating and collaborating also with groups and discussions (coordinator CAR is already part of AI4EU consortia and CAPRI consortium is already engaged into AI4EU open community).

5 CAPRI Reference Architecture

The Cognitive Automation Platform (CAP) Reference Architecture is defined leaning to the inputs of the D2.2 Use Case Requirements and the existing relevant reference architectures in the industrial domain already analysed in this document.

5.1 CAP Reference Architecture

The CAP Reference Architecture (RA) aims to support the entire data flow starting from the data collecting until the data consuming in particular:

- Support the acquisition of data from heterogeneous sources (IIoT and custom systems).
- Allow the historicization of data generated in the IoT or industrial field on ad-hoc storage
- Support the usage of data supporting applications.
- Provide a business analytics suite based on machine learning and cognitive algorithms.
- Allow the persistence of the output deriving from the analytics performed (both in terms of model and in terms of predictions).
- Manage business analytics based on batch data and streaming data.
- Allow the management of edge or wide scenarios.
- Integrate security modules for the user management.
- Support the Data Sovereignty principles.

---

[21] https://www.ai4eu.eu/forecasted-repository

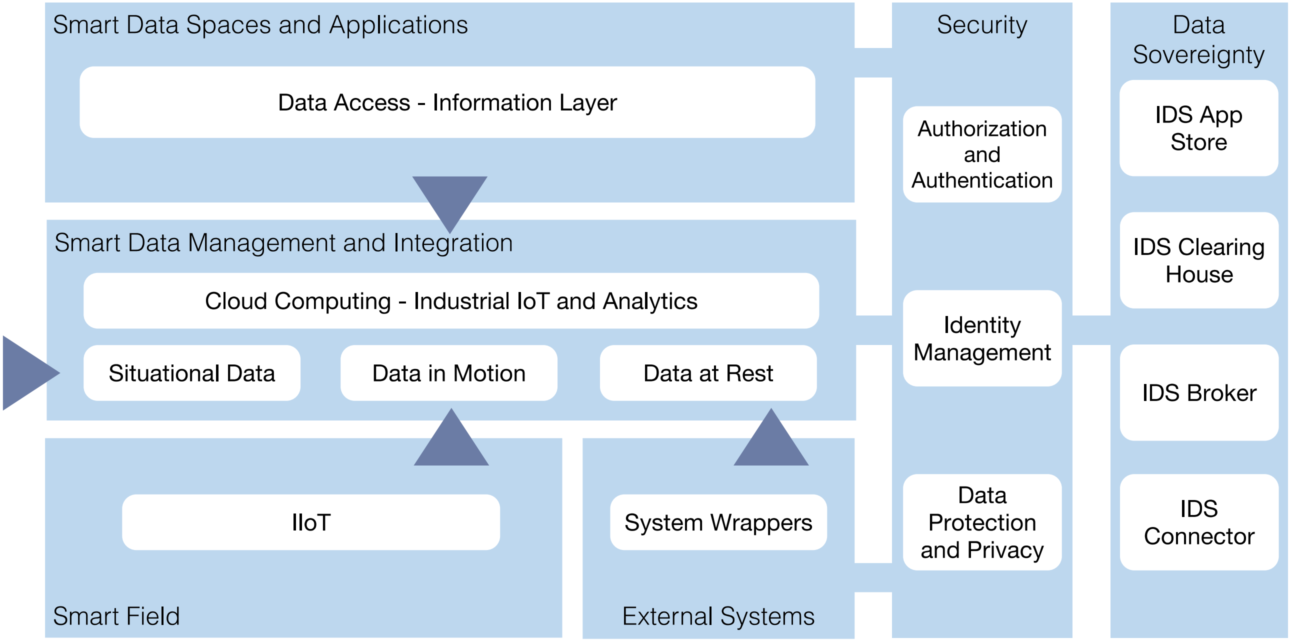


Figure 31: CAPRI Reference Architecture

The three-tier RA depicted in Figure 31 defines several functional macro-components:

- **Smart Field** contains the Industrial IoT (IIoT) physical layer, composed of machines, sensors, devices, actuators, etc and adapters.
- **External Systems** defines enterprise systems (ERPs, PLMs, customized, etc.) for the process supporting and adapters.
- **Smart Data Management and Integration** contains 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.
- **Smart Data Spaces and Applications** represents the data application services for representing and consuming historical, streaming and processed data.
- **Security** defines components for the authorization and authentication of users and systems. It integrates also modules for data protection and privacy.
- **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.

Due to the modularity of the architecture, it can be taken as a reference for describing both cloud and edge computing implementations. The Figure 32 shows the CAP RA for Edge paving the way for the data collection, storing, processing and presentation directly from the plant. In that sense, the CAP Reference Architecture ensure the capabilities to ingest, process and consume data with low latency and fast reactivity. Some cognitive services can be integrated and made available in the Edge Analytics module for supporting the operational intelligence and blue-collars activities in the plant. The Smart Data Management and Integration component will act as a data provider for the analogous cloud one in order to send edge data (data in motion and/or processed information) to the cloud for longer term processing and archiving.

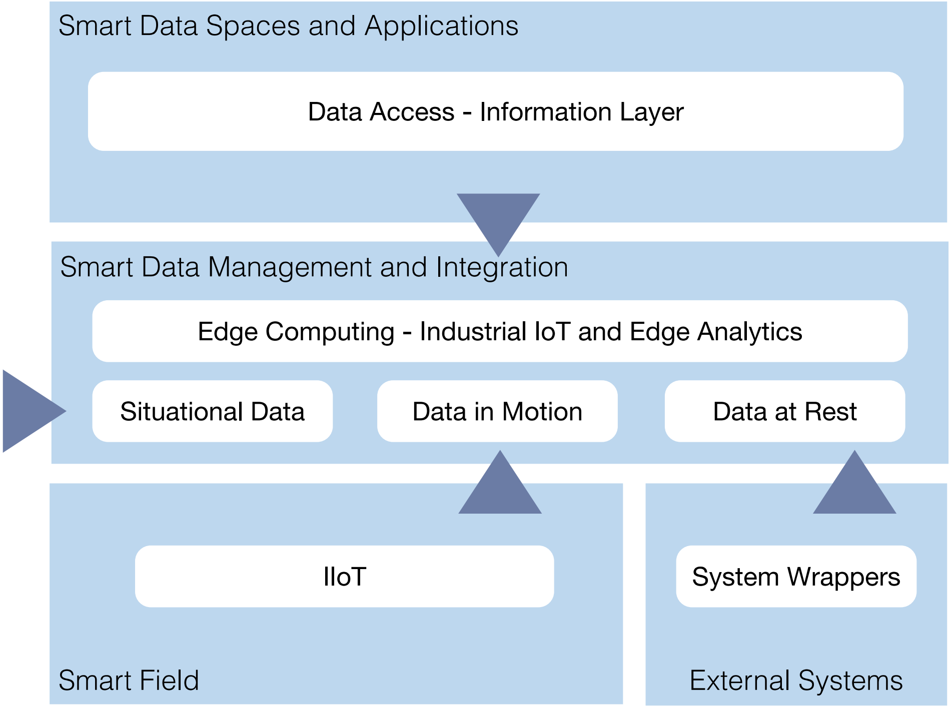


Figure 32: CAP Reference Architecture for Edge

5.1.1 Smart Field

The Smart Field component represents the physical layer and the interfaces for collecting data and communicating with IIoT systems. The Smart Field contains industrial devices, machines, actuators, sensors, wearable devices, etc. that are spread in the shop floor, and supports the most common industrial and, more in general, IoT protocols such as OPC UA, MQTT, etc. Standards interfaces and protocols must be used, in order to represent the information collected from the plant and to connect and integrate actuators for implementing the sensing and control mechanisms. Data will be collected typically as Data in Motion (DiM) since data coming from IIoT systems are dynamic and should be ingested and processed in real time.

5.1.2 External Systems

External Systems component contains all internal and IT systems for supporting industrial processes. It represents static information that comes from Legacy Systems and can be collected as Data at Rest (DaR). Custom interfaces and system wrappers are a crucial part of the component, aiming to share data using smart data models for representing information.

5.1.3 Smart Data Management and Integration

The Smart Data Management and Integration is the core of the architecture since it contains the brokering, the storage and the processing of the data.

Data are represented using standard information models and made available using standard APIs. Thought the service layer, data can be collected and persisted supporting a wide range of database (relational, time-series, etc.).

The data brokering component is used to acquire data from Smart Field and External Systems modules routing them to the correct destination. Data can be collected in different formats from different sources such as IoT sensors, social networks, etc. The ideal is to acquire large volumes of data with low latency. Communication between the components of the architecture is based on specific communication protocols that allow data collection. It is assumed that the pre-processing step has been performed to improve the quality of the data, for example through filtering operations to reduce redundancy and the amount of data to transferred.

The data storage module allows the storage of data during their life cycle within the Reference Architecture. The choice of the type of data storage components depends on the use case. For example, to archive the entire data history you need to rely on a scalable, fault-tolerant and high availability component. In the case of a data query, a component must be chosen so that it can provide an answer in the shortest possible time. The performance of traditional relational databases is not suitable for managing growing volumes of data and complex types of data. Hence, the widespread diffusion of NOSQL technologies.

The CAP Reference Architecture aims to cover several industrial scenarios, from the edge to the cloud processing, passing through the Big Data analytics. In that sense, supporting a wide spectrum of databases is not sufficient, advanced processing capabilities must be supported. The RA is able to support the analysis of streaming and batch data acquired from heterogeneous external sources with the support of machine learning technologies. In particular, a typical scenario in the IoT and industrial field is to react to events in real time based on the knowledge of past events, an essential information for predicting future behaviour, for example in order to identify any anomalies. In that sense, the platform integrates cognitive computing services in order to learn from experience and derive insights to unlock the value of big data.

5.1.4 Smart Data Spaces and Applications

The CAP RA supports the creation and integration of smart data applications. The Smart Data Spaces and Applications, in fact, contains the system and user applications for presenting and consuming data. A wide spectrum of domains and class of applications are supported:

- BI & Analytics increasing the business value supporting augmented intelligence and machine learning for implementing data-driven and cognitive decision-making.
- AR/VR offering services for supporting the decision-making process (improve the human-machine interactions, accomplish proficient operational intelligence, etc.).
- Chatbots & Virtual Assistants to enrich (chatbots) and assists (virtual assistants) the users (blue-collars and customers).
- Novel HMI implementing new user experience developments such as supporting inexperienced operators, machine-human-machine operations, operator decision-making, etc.
- Self-service Visualization supporting business users for accessing all data features and make data-driven decisions in a quick and scalable way.
- Generic Cognitive Applications, implementing the cognitive manufacturing such as the self-learning, the continuous learning, the machine reasoning, the communication in natural language, etc.

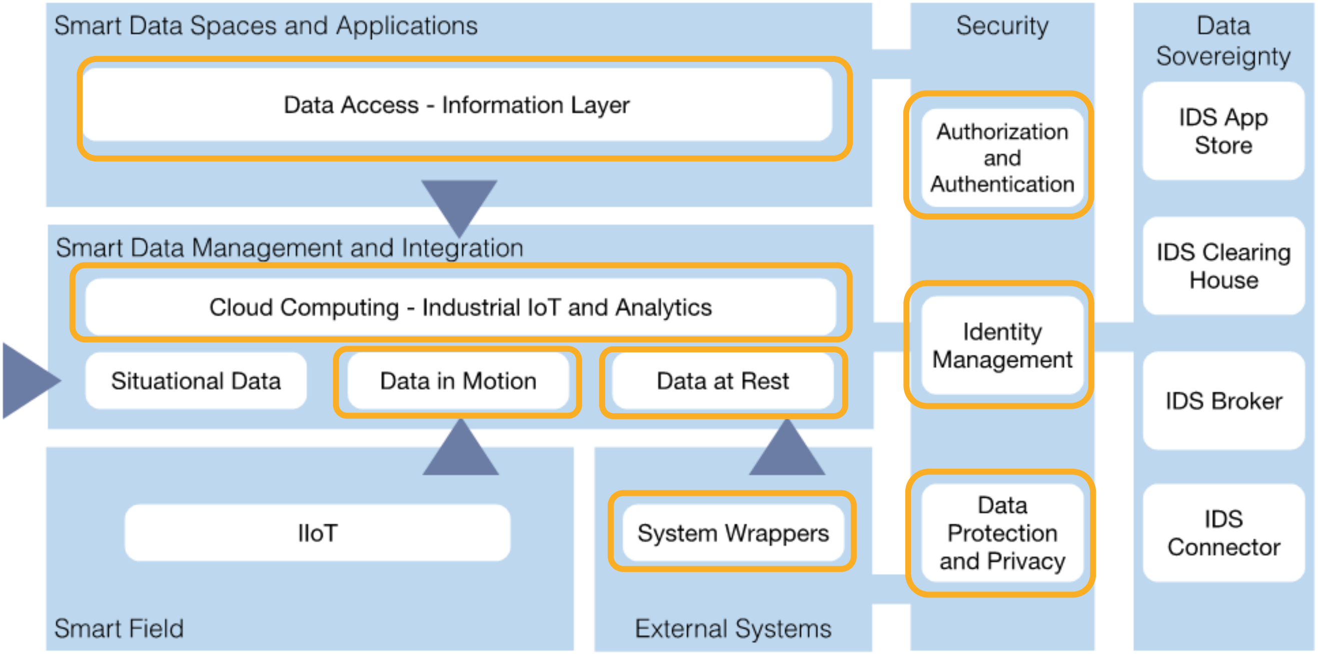
5.2 Mapping of the CAPRI Cognitive Solution

This sub-chapter defines the mapping between each cognitive solution and the Reference Architecture described above. In such a context, each cognitive solution will be evaluated at the four levels from sensors to planning, and for each Case Study (Asphalt, Steel and Pharma), as well as checked against the defined CAPRI RA in order to assess which are the most valuable components to deliver cognition capabilities in the particular scope of the solution and use case.

To this end, this section focuses on highlighting the particular aspects of the CAPRI RA that impact the design, development and deploy of the cognitive solutions in the relative Case Studies.

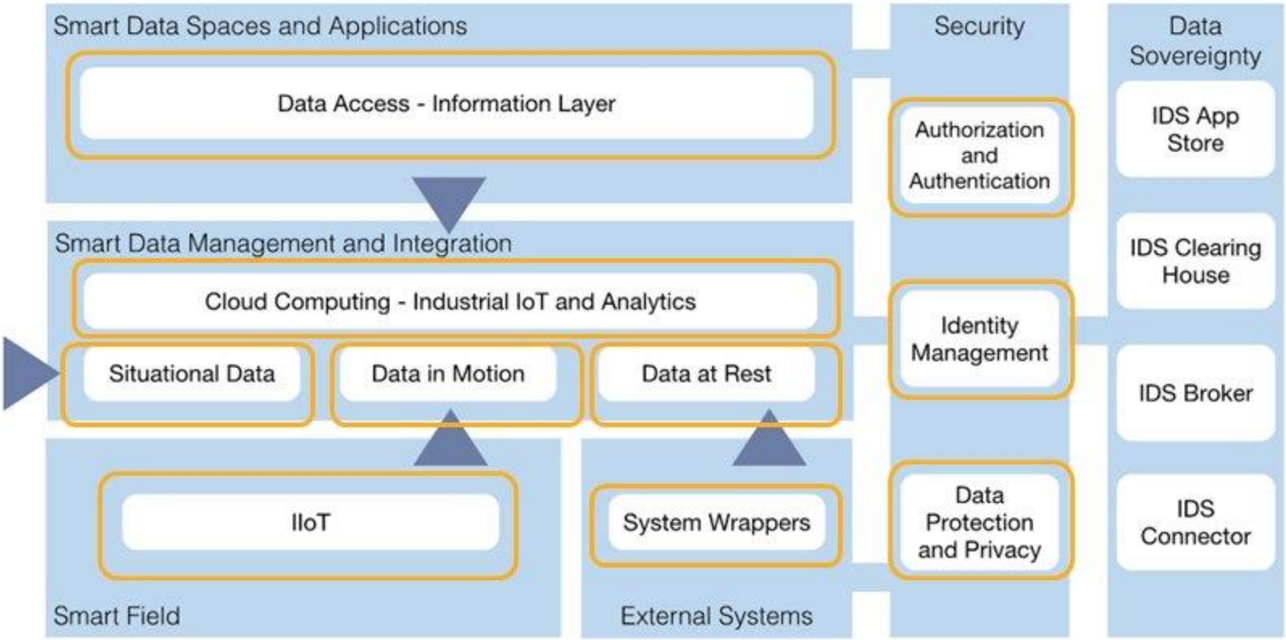
5.2.1 Asphalt

5.2.1.1 (CAC1) COGNITIVE CONTROL OF DRYER DRUM



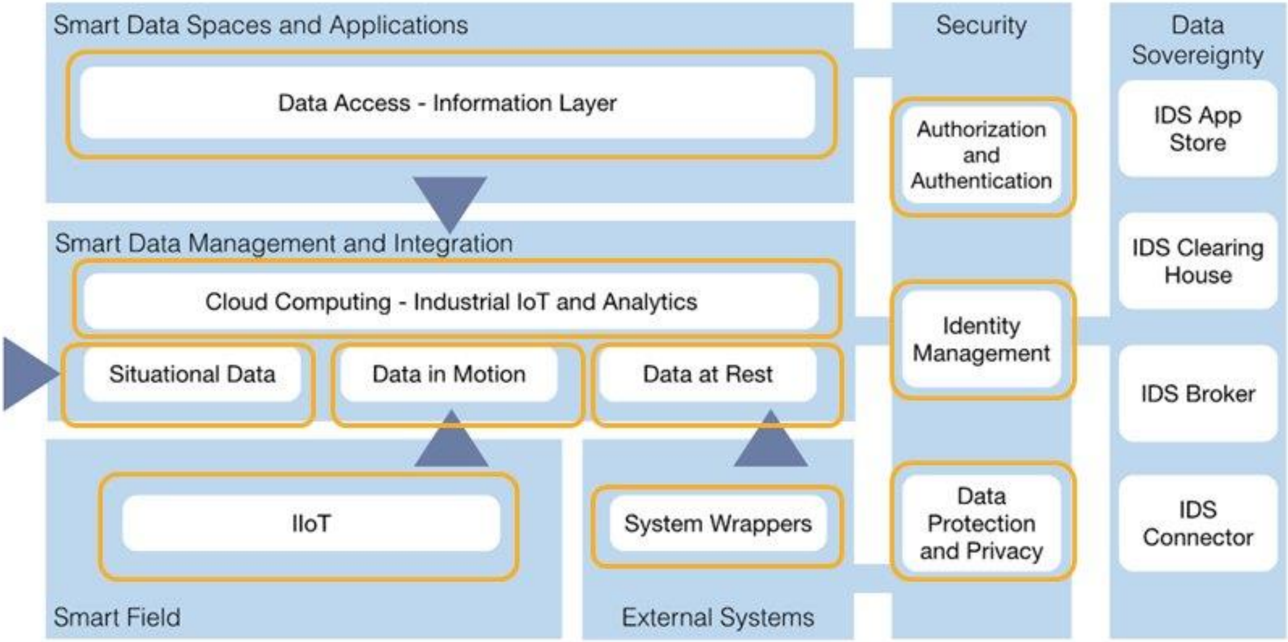
- ✓ **IIoT –** collection of data coming from sensors and actuators
- ✓ **Data in Motion –** real time generation of different setpoints of the temperature of the hot aggregates coming out of the drum and the temperature of the combustion gases leaving the dryer and entering the filter.
- ✓ **Situational Data** – management and aggregation of context data
- ✓ **Smart Data Space and Application –** cognitive algorithm for the control of dryer drum (according with real-time constraints this could be provided through edge nodes instead of the cloud)
- ✓ **Data at rest –** management of static data coming from internal and external sources
- ✓ **System Wrappers** – historical data coming from SCADA system can be acquired
- ✓ **Authorization and Authentication** – only authorized personnel can access the system
- ✓ **Identity Management** – only identified personnel can access the system
- ✓ **Data Protection and Privacy –** data needs to be protected (industrial property and for GDPR compliance)

5.2.1.2 (CAO1) PREDICTIVE MAINTENANCE OF BAGHOUSE BASED ON COGNITIVE SENSORS AND EXPERT KNOWLEDGE



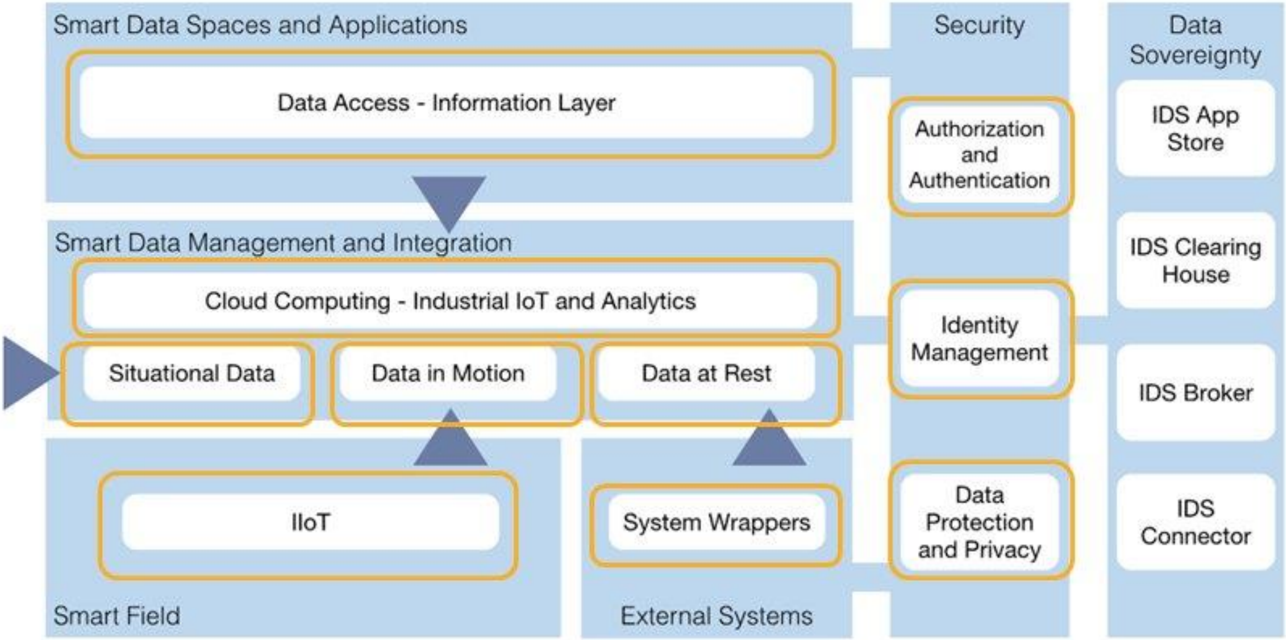
- ✓ **IIoT** – collection of data coming from sensors and actuators
- ✓ **Data in Motion** –calculation of pressure between the input and output to determine dust accumulation, and need to clean up filters. Alarms can be generated and showcased to the operators
- ✓ **Situational Data** – management and aggregation of context data (pressure, temperature, humidity, etc.)
- ✓ **Smart Data Space and Application** – cognitive algorithm for the predictive maintenance of the baghouse filter (according with constraints this could be provided through edge nodes instead of the cloud)
- ✓ **Data at rest** – management of static data coming from internal and external sources
- ✓ **System Wrappers** – historical data coming from SCADA system can be acquired (orders, historical data, etc.)
- ✓ **Authorization and Authentication** – only authorized personnel can access the system
- ✓ **Identity Management** – only identified personnel can access the system
- ✓ **Data Protection and Privacy** – data needs to be protected (industrial property and for GDPR compliance)

5.2.1.3 (CAS1) COGNITIVE SENSOR OF BITUMEN CONTENT IN RECYCLED ASPHALT



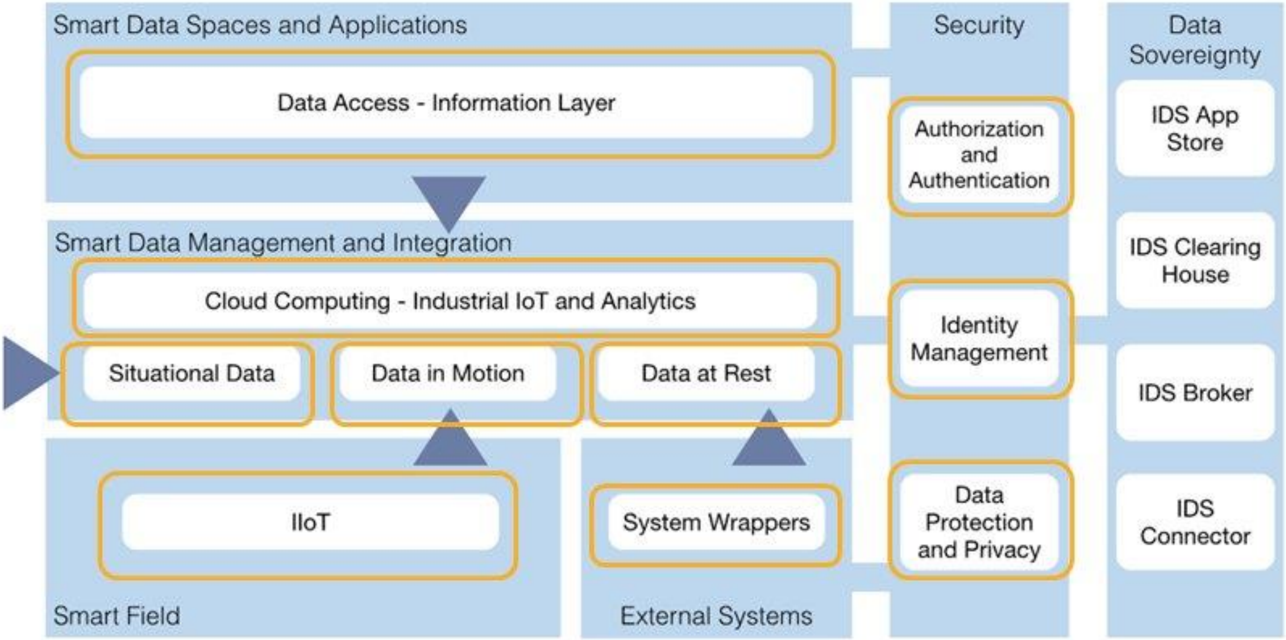
- ✓ **IIoT** – collection of raw data coming from multispectral detectors and illumination sources
- ✓ **Data in Motion** – real time calculation of RAP images and extraction of critical characteristic to measure the amount of bitumen
- ✓ **Situational Data** – management and aggregation of context data (temperature, humidity, etc.)
- ✓ **Smart Data Space and Application** – cognitive algorithm for the optimization of the asphalt mix formula (this will be run in a custom made hardware provided by AIMEN).
- ✓ **Data at rest** – management of static data coming from internal and external sources
- ✓ **System Wrappers** – historical data coming from SCADA system can be acquired (historical data, etc.)
- ✓ **Authorization and Authentication** – only authorized personnel can access the system
- ✓ **Identity Management** – only identified personnel can access the system
- ✓ **Data Protection and Privacy** – data needs to be protected (industrial property and for GDPR compliance)

5.2.1.4 (CAS2) COGNITIVE SENSOR FOR PARTICLE SIZE MEASUREMENT



- ✓ **IIoT** – collection of data coming from sensors (filler flow entering in the baghouse, weight of filler that exits the baghouse)
- ✓ **Data in Motion** – real time processing of operational data to measure the fine filler particle size quantity that exits to the baghouse filter and that one that is poured on to the hot bins and generate the settings of the setpoints of the controllers.
- ✓ **Situational Data** – management and aggregation of context data (% filler in the final mix, etc.)
- ✓ **Smart Data Space and Application** – cognitive algorithm for the extraction of knowledge to correct the amount of material to be fed in the process, and optimize the overall production system (according with real-time constraints this could be provided through edge nodes instead of the cloud)
- ✓ **Data at rest** – management of static data coming from internal and external sources
- ✓ **System Wrappers** – historical data coming from external systems can be acquired (historical data, experimental curve laboratory, etc.)
- ✓ **Authorization and Authentication** – only authorized personnel can access the system
- ✓ **Identity Management** – only identified personnel can access the system
- ✓ **Data Protection and Privacy** – data needs to be protected (industrial property and for GDPR compliance)

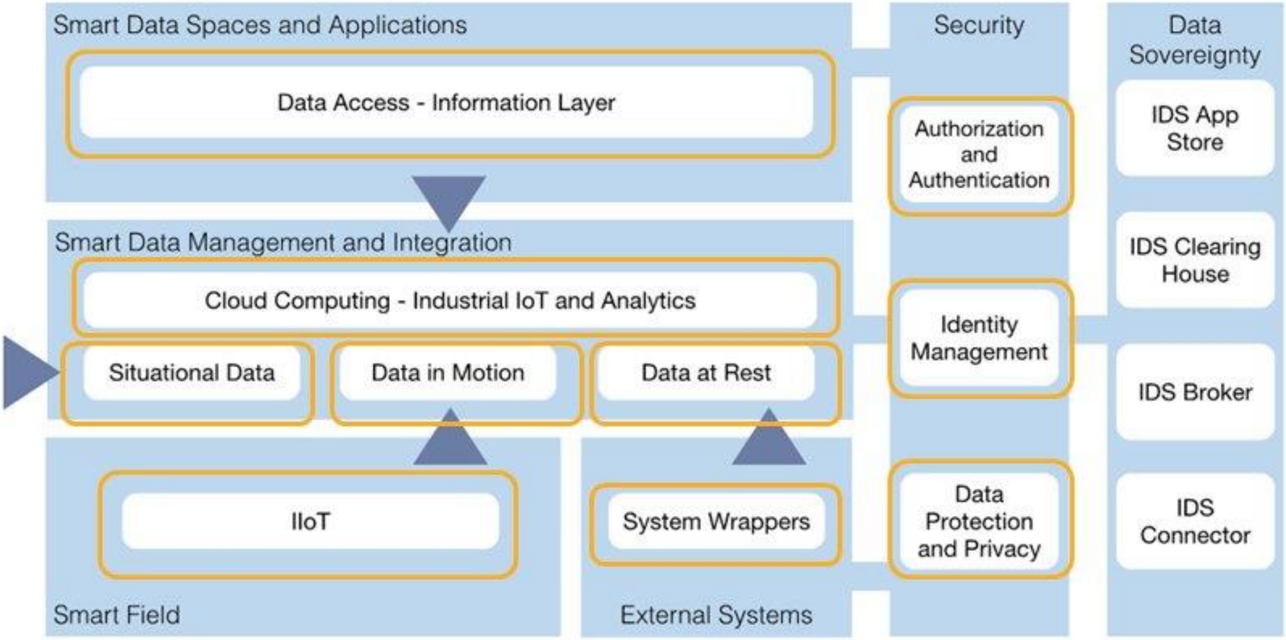
5.2.1.5 (CAP1) COGNITIVE SOLUTION OF PLANNING AND CONTROL OF ASPHALT PRODUCTION



- ✓ **IIoT** – collection of data coming from various cognitive sensors, or other sensors
- ✓ **Data in Motion** – real time processing of operational data with adjustment to generate setpoints of total flow to dry and % of the cold aggregates to optimize production. Alarms can be generated and showcased to the operators
- ✓ **Situational Data** – management and aggregation of context data (pressure, temperature, humidity, etc.)
- ✓ **Smart Data Space and Application** – cognitive algorithm for the generation of the settings of the set points of the controllers, the amounts of material to be incorporated into the processes such as the mixer and other data that will allow to optimize the production process (according with real-time constraints this could be provided through edge nodes instead of the cloud). Cognitive algorithms to decide the best moment to start production so that asphalt arrives in the best possible conditions to the place where it will be applied
- ✓ **Data at rest** – management of static data coming from internal and external sources
- ✓ **System Wrappers** – historical data coming from SCADA system can be acquired (historical data, etc.)
- ✓ **Authorization and Authentication** – only authorized personnel can access the system
- ✓ **Identity Management** – only identified personnel can access the system
- ✓ **Data Protection and Privacy** – data needs to be protected (industrial property and for GDPR compliance)

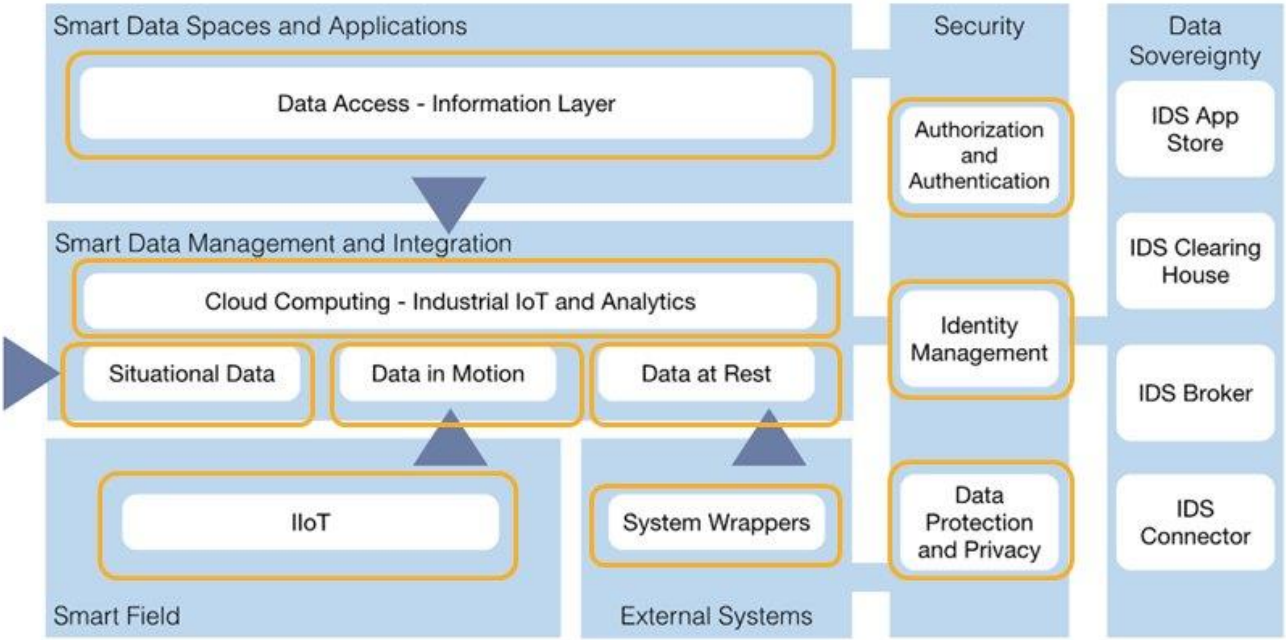
5.2.2 Steel

5.2.2.1 (CSS1) COGNITIVE STEEL TRACKING SENSOR



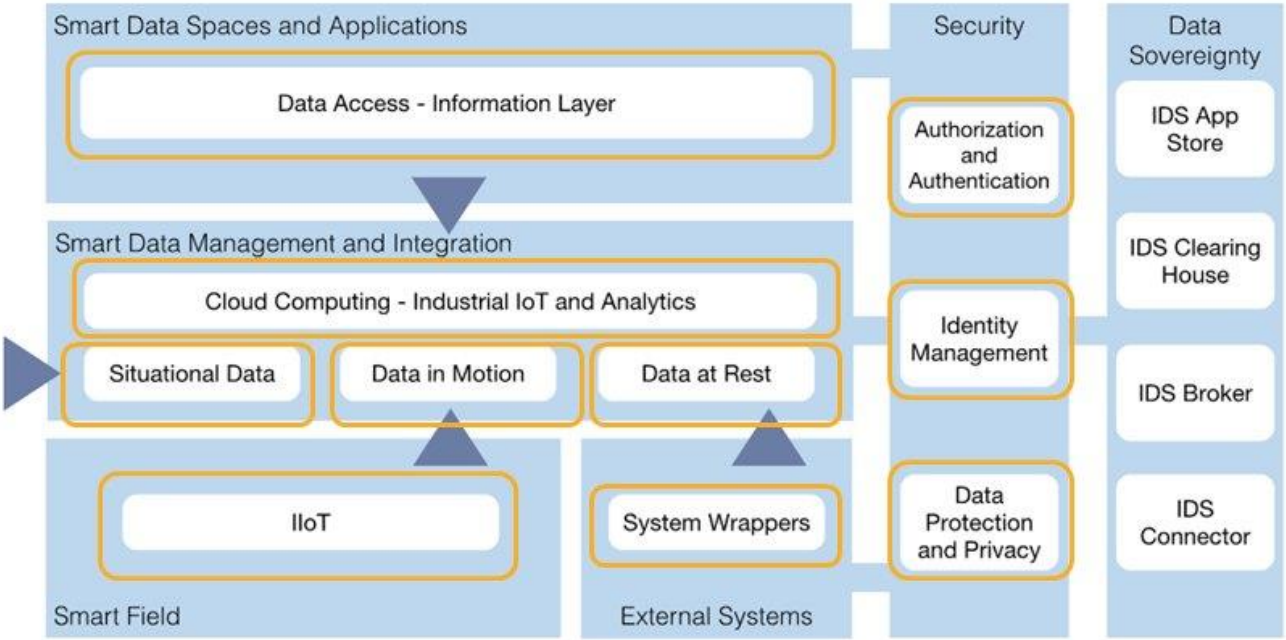
- ✓ **IIoT –** collection of data coming from camera scanners and laser engraving systems to feed further processing
- ✓ **Data in Motion** – real time processing of operational data within the casting machine, rolling mill and finishing line
- ✓ **Situational Data** – management and aggregation of context data (location, temperature, etc.)
- ✓ **Smart Data Space and Application** –in the casting machine billets are marked with a QR code that identifies the heat they belong to, then in the rolling mill a process tracking and a marking of the hot bars related to the production order is done. The bars are again identified in the finishing line. (according with real-time constraints this could be provided through edge nodes instead of the cloud)
- ✓ **Data at rest** – management of static data coming from internal and external sources
- ✓ **System Wrappers** – historical data coming from MES and SCADA system can be acquired (historical data, etc.)
- ✓ **Authorization and Authentication** – only authorized personnel can access the system
- ✓ **Identity Management** – only identified personnel can access the system
- ✓ **Data Protection and Privacy** – data needs to be protected (industrial property)

5.2.2.2 (CSS2) COGNITIVE SOLIDIFICATION SENSOR



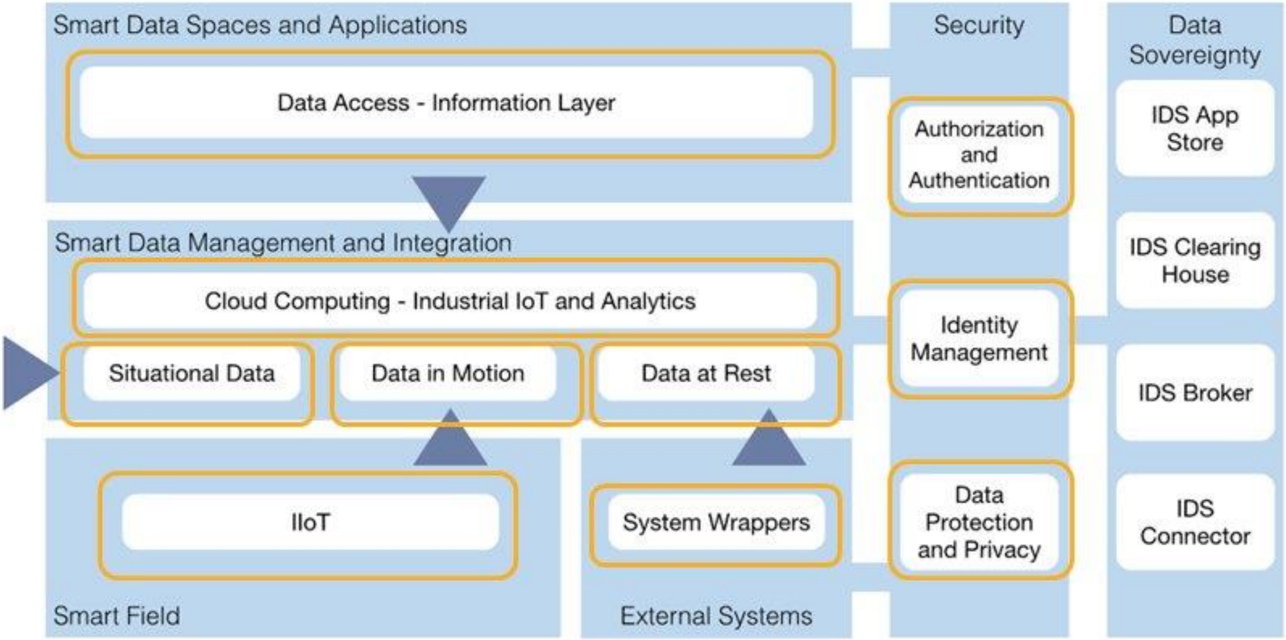
- ✓ **IIoT –** collection of data coming from various sensors
- ✓ **Data in Motion –** real time processing of temperature, spray water flow, etc. in the casting process
- ✓ **Situational Data –** management and aggregation of context data (temperature, etc.)
- ✓ **Smart Data Space and Application –** simulation algorithm for the tracking and monitoring of the casting process, with supervision of the temperature profile and solidification front in the casting strands (according with real-time constraints this could be provided through edge nodes instead of the cloud)
- ✓ **Data at rest –** management of static data coming from internal and external sources. Possibility to perform replay simulations on historical data.
- ✓ **System Wrappers –** historical data coming from MES and SCADA system can be acquired (historical data, etc.)
- ✓ **Authorization and Authentication –** only authorized personnel can access the system
- ✓ **Identity Management –** only identified personnel can access the system
- ✓ **Data Protection and Privacy –** data needs to be protected (industrial property)

5.2.2.3 (CSS3) COGNITIVE TEMPERATURE SENSOR



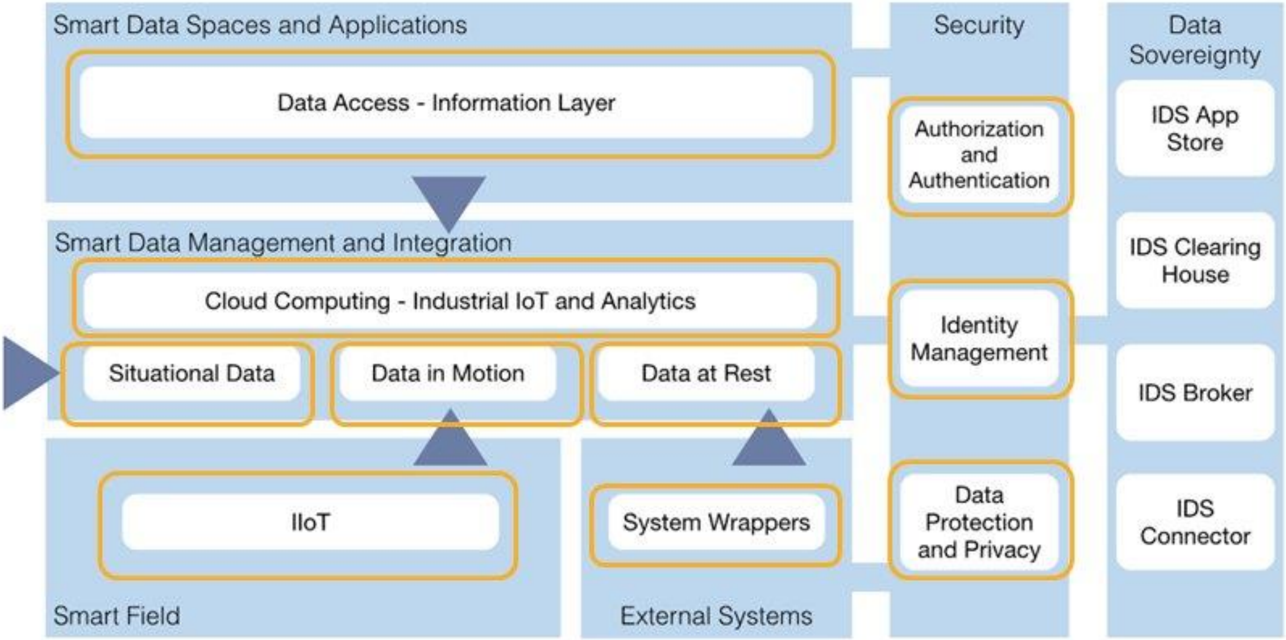
- ✓ **IIoT** – collection of data coming from the cognitive solidification sensor and other sensors
- ✓ **Data in Motion** – real time processing of temperature and position data of the product in production
- ✓ **Situational Data** – management and aggregation of context data (pressure, temperature, etc.)
- ✓ **Smart Data Space and Application** – cognitive algorithm for the continuous monitoring of the temperature state of the product (according with real-time constraints this could be provided through edge nodes instead of the cloud)
- ✓ **Data at rest** – management of static data coming from internal and external sources. Possibility to perform replay simulations on historical data.
- ✓ **System Wrappers** – historical data coming from MES and SCADA system can be acquired (historical temperature profiles, etc.)
- ✓ **Authorization and Authentication** – only authorized personnel can access the system
- ✓ **Identity Management** – only identified personnel can access the system
- ✓ **Data Protection and Privacy** – data needs to be protected (industrial property)

5.2.2.4 (CSS4) COGNITIVE SCALE SENSOR



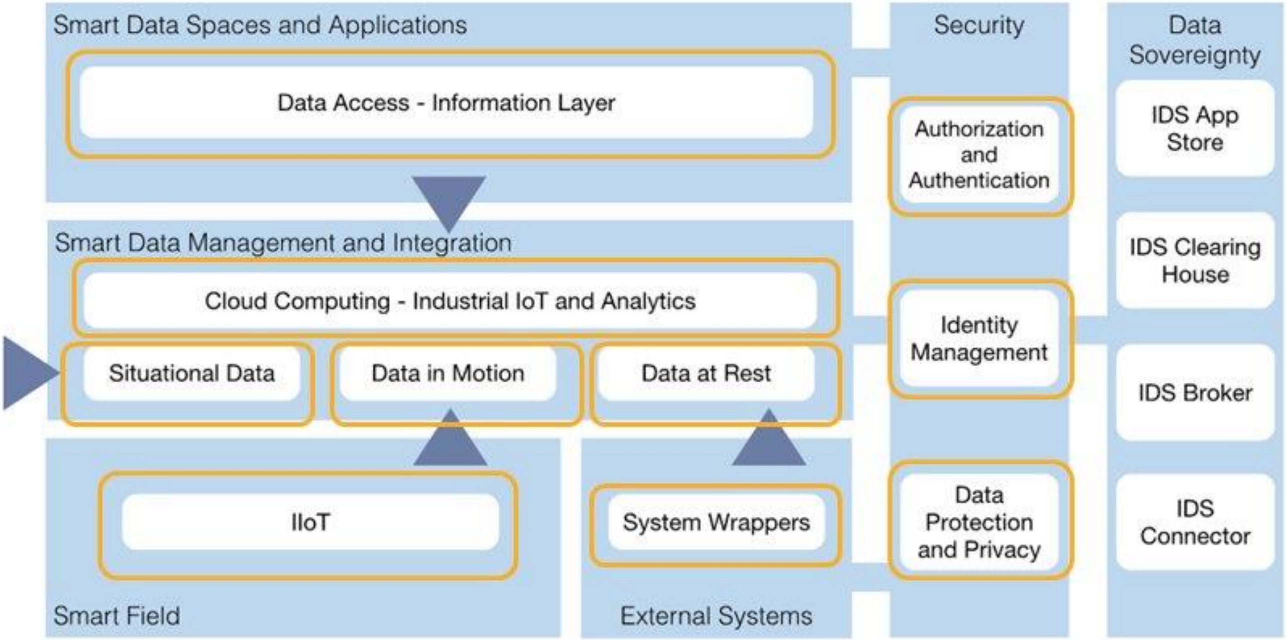
- ✓ **IIoT** – collection of data coming from the cognitive temperature sensor and other sensors
- ✓ **Data in Motion** – real time processing of temperature and position data of the product in production
- ✓ **Situational Data** – management and aggregation of context data (pressure, humidity, temperature, other chemical properties, etc.)
- ✓ **Smart Data Space and Application** – cognitive algorithm will predict the amount of scale on the surface after the rolling processes in terms of chemical composition and thickness of the scale layer (according with real-time constraints this could be provided through edge nodes instead of the cloud)
- ✓ **Data at rest** – management of static data coming from internal and external sources Possibility to perform replay simulations on historical data.
- ✓ **System Wrappers** – historical data coming from Digital Twin system can be acquired
- ✓ **Authorization and Authentication** – only authorized personnel can access the system
- ✓ **Identity Management** – only identified personnel can access the system
- ✓ **Data Protection and Privacy** – data needs to be protected (industrial property)

5.2.2.5 (CSS5) COGNITIVE RISK AND ANOMALIES SENSOR



- ✓ **IIoT** – collection of data coming from various cognitive sensors, or other sensors
- ✓ **Data in Motion** – real time monitoring of the enhanced product data
- ✓ **Situational Data** – management and aggregation of context data (pressure, temperature, etc.)
- ✓ **Smart Data Space and Application** – cognitive algorithm based on the analysis of the process data stream to identify process anomalies, to prevent potential damages, steer processes back into the normal operation mode, and helping to repair the product by adapting the upcoming process chain.
- ✓ **Data at rest** – management of static data coming from internal and external sources Possibility to perform replay simulations on historical data.
- ✓ **System Wrappers** – historical data coming from MES and SCADA system can be acquired
- ✓ **Authorization and Authentication** – only authorized personnel can access the system
- ✓ **Identity Management** – only identified personnel can access the system
- ✓ **Data Protection and Privacy** – data needs to be protected (industrial property)

5.2.2.6 DIGITAL TWIN

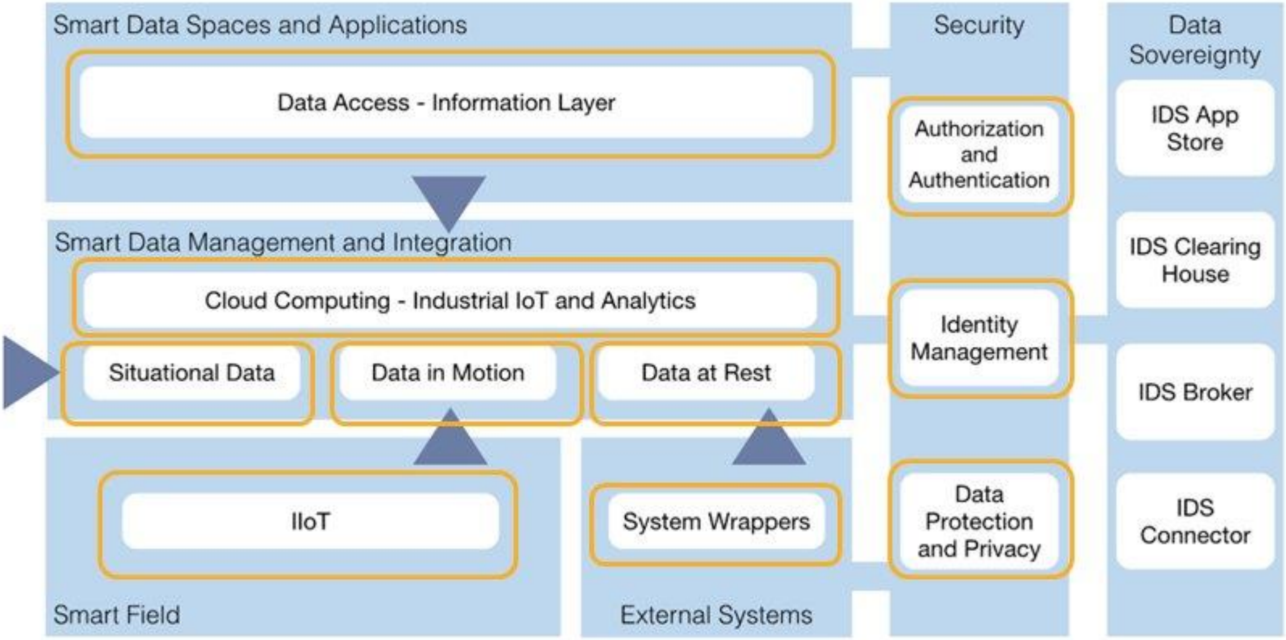


- ✓ **IIoT** – collection of data coming from various cognitive sensors and other sensors
- ✓ **Data in Motion** – real time processing of operational data
- ✓ **Situational Data** – management and aggregation of context data (temperature, humidity, etc.)
- ✓ **Smart Data Space and Application** – the digital twin approach is realized for products and machinery. The twin does not only represent a simulation but actively accompanies the product through the whole chain. The data of the digital twin encompasses first-principle models, hybrid models combining physical models with historical data and full data driven approaches. The whole data collected from CSS1-CSS5 will foster the machine learning procedures. Using the labels, supervised machine learning algorithms which are the backbone of the models for the cognitive sensors, can be pre-trained to provide data-driven models across the chain. Unsupervised machine learning algorithms can be continuously trained during production, building upon the supervised pre-training
- ✓ **Data at rest** – management of static data coming from internal and external sources Possibility to perform replay simulations on historical data.
- ✓ **System Wrappers** – historical data coming from MES and SCADA system can be acquired
- ✓ **Authorization and Authentication** – only authorized personnel can access the system
- ✓ **Identity Management** – only identified personnel can access the system
- ✓ **Data Protection and Privacy** – data needs to be protected (industrial property and potentially for GDPR compliance)

This project receives funding in the European Commission's Horizon 2020 Research Programme under Grant Agreement Number 870062

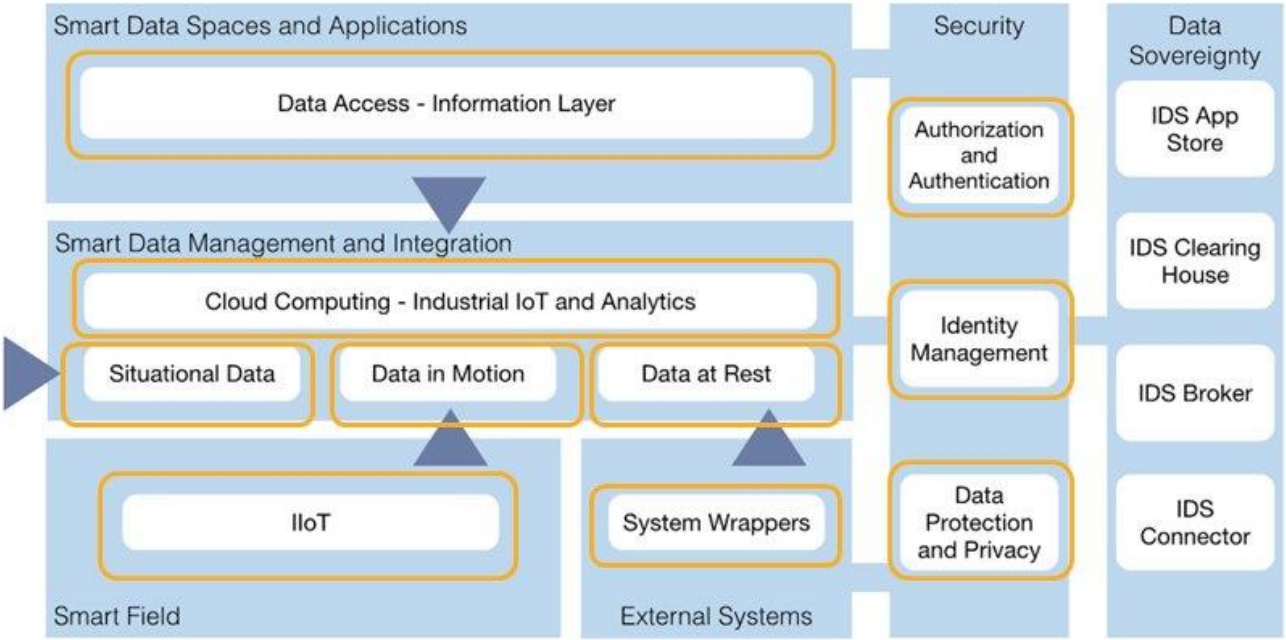
5.2.3 Pharma

5.2.3.1 (CPS1) BLEND UNIFORMITY



- ✓ **IIoT** – collection of data coming from various sensors and other computing nodes (Consigma, Raman, SIPAT)
- ✓ **Data in Motion** – real time processing of operational data coming from OPC server
- ✓ **Situational Data –** management and aggregation of context data (Temperature, powder mass flow, liquid mass flow, PhAT Probe, etc.)
- ✓ **Smart Data Space and Application** – cognitive algorithm computing the API concentration from Raman raw spectra (according with real-time constraints this could be provided through edge nodes instead of the cloud).
- ✓ **Data at rest** – management of static data coming from internal and external sources
- ✓ **System Wrappers** – historical data coming from PLC and SCADA system can be acquired
- ✓ **Authorization and Authentication** – only authorized personnel can access the system
- ✓ **Identity Management** – only identified personnel can access the system
- ✓ **Data Protection and Privacy –** data needs to be protected (industrial property and for GDPR compliance)

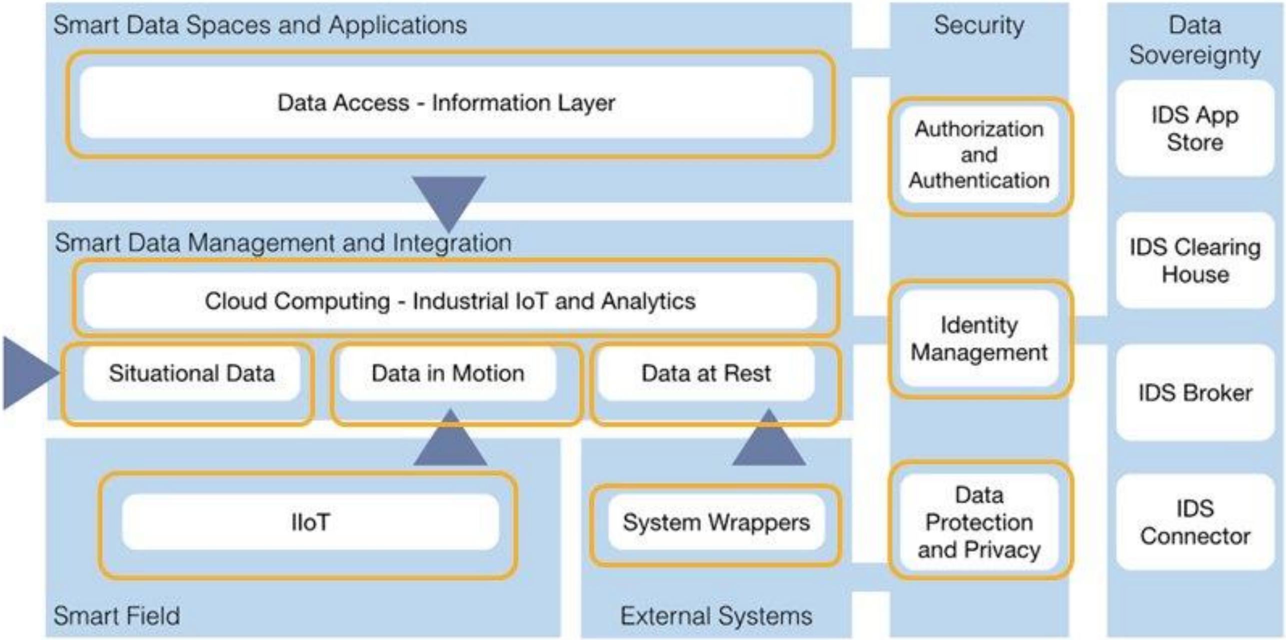
5.2.3.2 (CPS2) GRANULE QUALITY



- ✓ **IIoT** – collection of data coming from various sensors and other computing nodes (Consigma, Parsum, SIPAT)
- ✓ **Data in Motion** – real time processing of operational data coming from OPC server
- ✓ **Situational Data** – management and aggregation of context data (Temperature, powder mass flow, liquid mass flow, etc.)
- ✓ **Smart Data Space and Application** – cognitive algorithm computing a characteristic granule size descriptor from particle size raw data (according with real-time constraints this could be provided through edge nodes instead of the cloud).
- ✓ **Data at rest** – management of static data coming from internal and external sources
- ✓ **System Wrappers** – historical data coming from PLC and SCADA system can be acquired
- ✓ **Authorization and Authentication** – only authorized personnel can access the system
- ✓ **Identity Management** – only identified personnel can access the system
- ✓ **Data Protection and Privacy** – data needs to be protected (industrial property and for GDPR compliance)

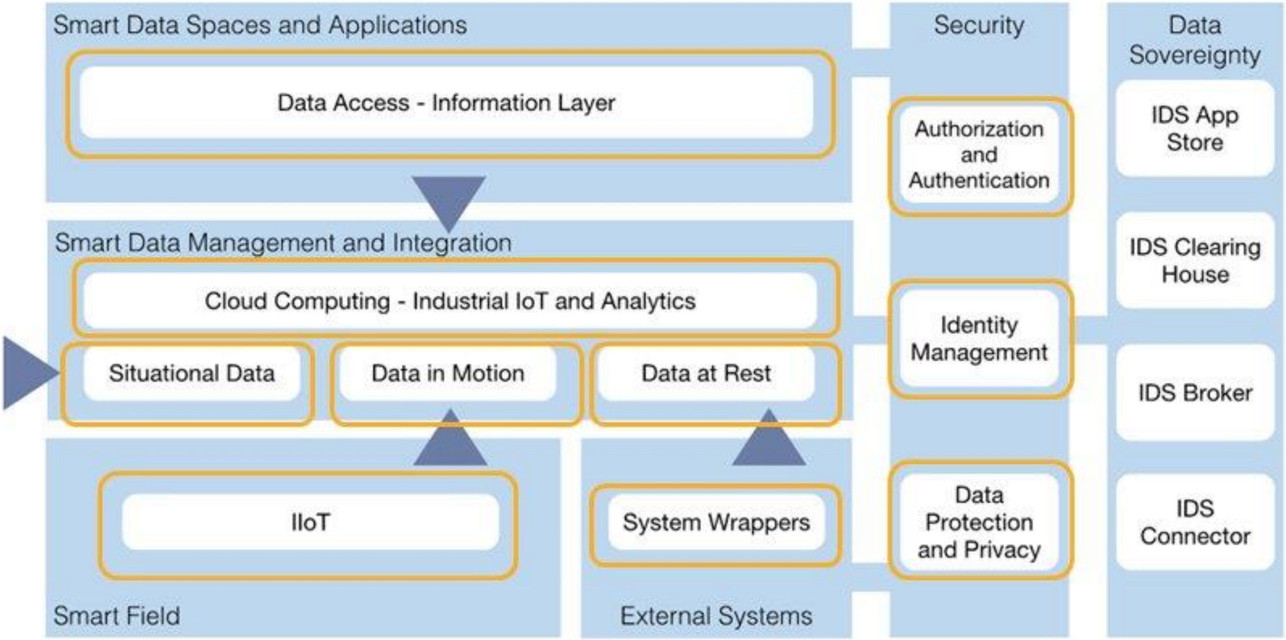
This project receives funding in the European Commission's Horizon 2020 Research Programme under Grant Agreement Number 870062

5.2.3.3 (CPS3) GRANULE MOISTURE



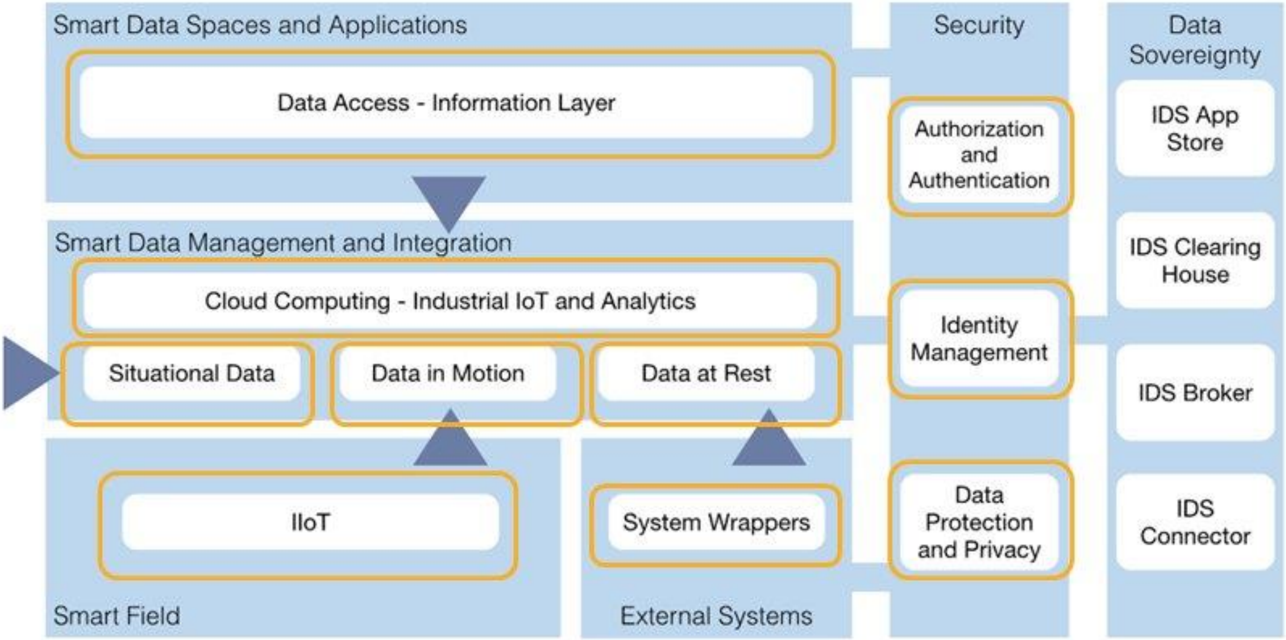
- ✓ **IIoT** – collection of data coming from various sensors and other computing nodes (Consigma, Parsum, SIPAT)
- ✓ **Data in Motion** – real time processing of operational data coming from OPC server
- ✓ **Situational Data** – management and aggregation of context data (Temperature, powder mass flow, liquid mass flow, etc.)
- ✓ **Smart Data Space and Application** – cognitive algorithm based on a process model of the dryer and available process data, computing the granule moisture in the dryer (according with real-time constraints this could be provided through edge nodes instead of the cloud).
- ✓ **Data at rest** – management of static data coming from internal and external sources
- ✓ **System Wrappers** – historical data coming from PLC and SCADA system can be acquired
- ✓ **Authorization and Authentication** – only authorized personnel can access the system
- ✓ **Identity Management** – only identified personnel can access the system
- ✓ **Data Protection and Privacy** – data needs to be protected (industrial property and for GDPR compliance)

5.2.3.4 (CPS4) DISSOLUTION PREDICTION



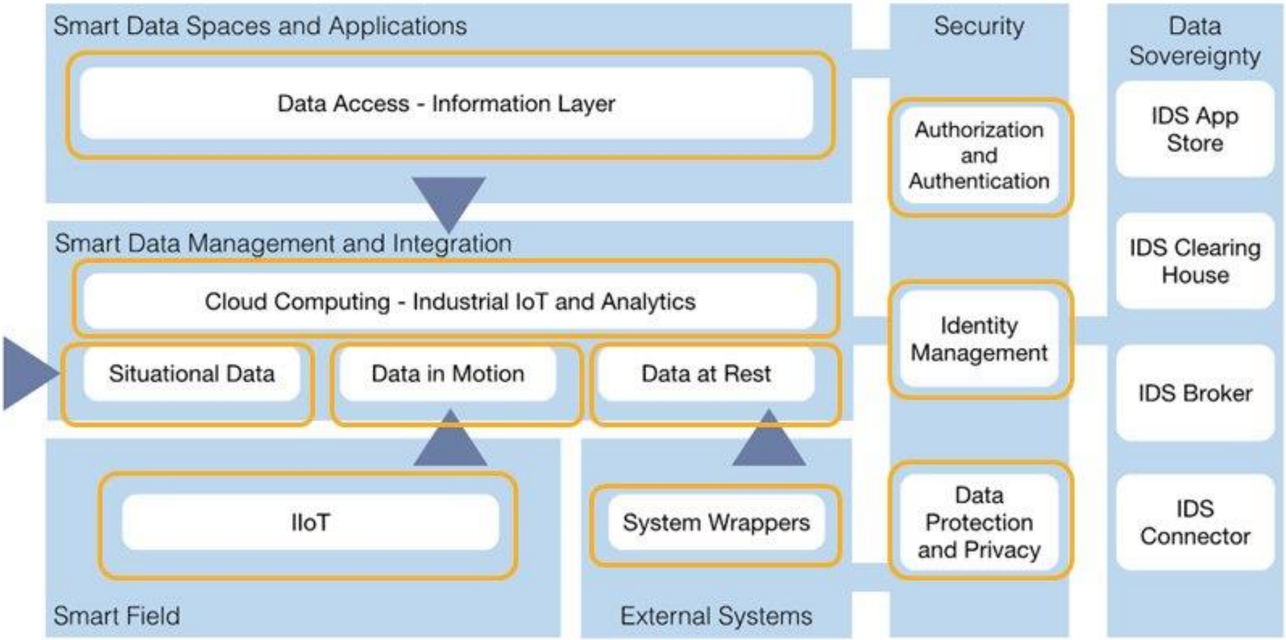
- ✓ **IIoT –** collection of data coming from various sensors, other computing nodes (Consigma, Parsum, Raman, SIPAT) and cognitive (CPS1-CPS3) sensors.
- ✓ **Data in Motion –** real time processing of operational data coming from OPC server
- ✓ **Situational Data** – management and aggregation of context data (Temperature, powder mass flow, liquid mass flow, etc.)
- ✓ **Smart Data Space and Application** – cognitive algorithm for dissolution prediction based on a data driven model capable of predicting dissolution profiles from process data.
- ✓ **Data at rest** – management of static data coming from internal and external sources
- ✓ **System Wrappers** – historical data coming from PLC and SCADA system can be acquired
- ✓ **Authorization and Authentication** – only authorized personnel can access the system
- ✓ **Identity Management** – only identified personnel can access the system
- ✓ **Data Protection and Privacy** – data needs to be protected (industrial property and for GDPR compliance)

5.2.3.5 (CPS5) FAULT DETECTION



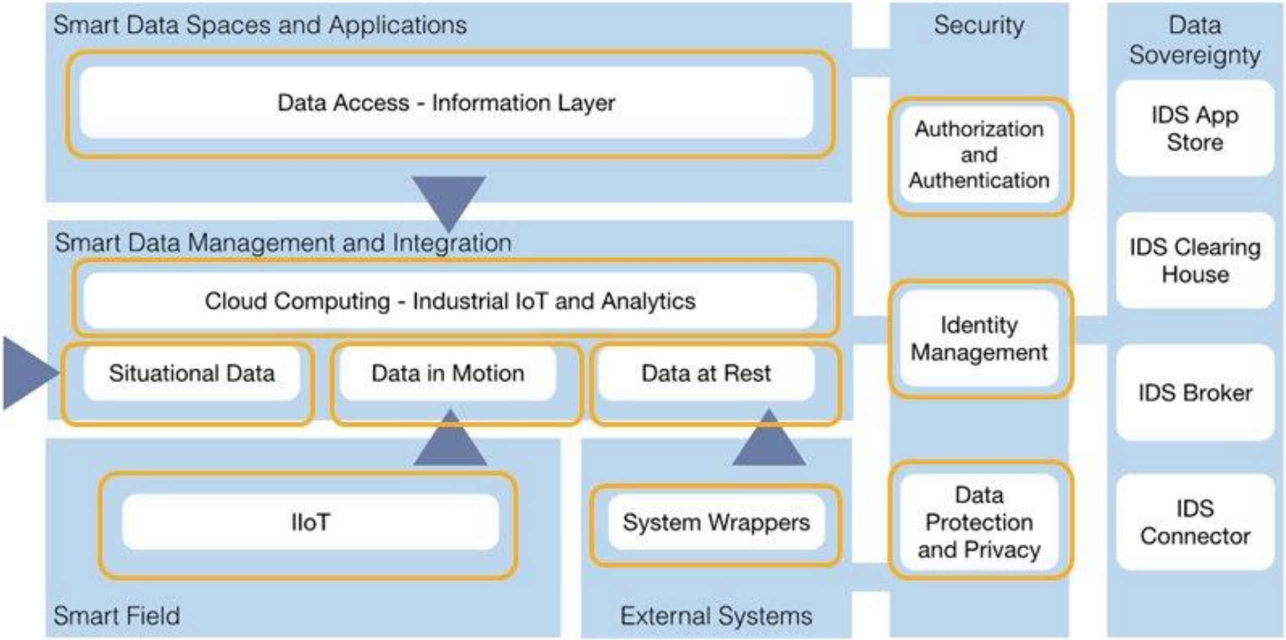
- ✓ **IIoT** – collection of data coming from various sensors, other computing nodes (Consigma, Parsum, Raman, SIPAT) and cognitive (CPS1-CPS3) sensors
- ✓ **Data in Motion** – real time processing of operational data coming from OPC server
- ✓ **Situational Data** – management and aggregation of context data (Temperature, powder mass flow, liquid mass flow, etc.)
- ✓ **Smart Data Space and Application** – cognitive algorithm for fault detection based on process data (according with real-time constraints this could be provided through edge nodes instead of the cloud).
- ✓ **Data at rest** – management of static data coming from internal and external sources
- ✓ **System Wrappers** – historical data coming from PLC and SCADA system can be acquired
- ✓ **Authorization and Authentication** – only authorized personnel can access the system
- ✓ **Identity Management** – only identified personnel can access the system
- ✓ **Data Protection and Privacy** – data needs to be protected (industrial property and for GDPR compliance)

5.2.3.6 (CPC1) COGNITIVE CONTROL CONCEPT



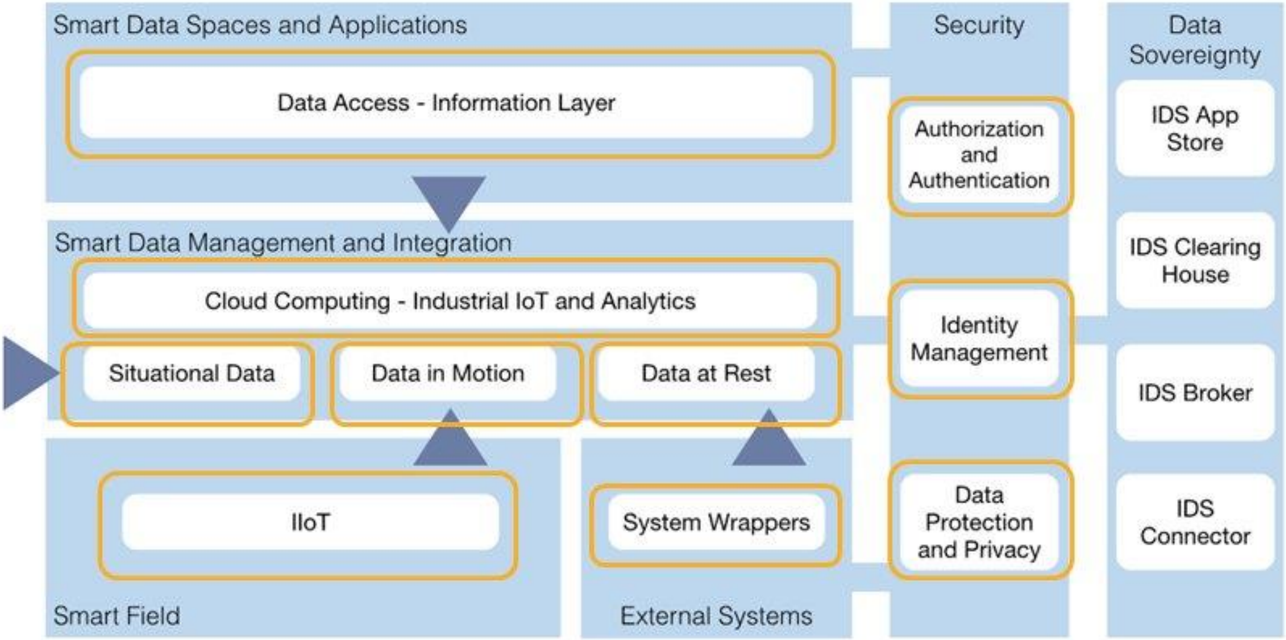
- ✓ **IIoT** – collection of data coming from various sensors, actuators and cognitive (CPS1-CPS4) sensors.
- ✓ **Data in Motion** – real time processing of operational data coming from OPC server
- ✓ **Situational Data** – management and aggregation of context data (particle size, granule moisture, drying settings, etc.)
- ✓ **Smart Data Space and Application** – cognitive control algorithm based on an MPC control block, with model predictive control of particle size and granule moisture, and a model based discharge strategy (according with real-time constraints this could be provided through edge nodes instead of the cloud).
- ✓ **Data at rest** – management of static data coming from internal and external sources
- ✓ **System Wrappers** – historical data coming from PLC and SCADA system can be acquired
- ✓ **Authorization and Authentication** – only authorized personnel can access the system
- ✓ **Identity Management** – only identified personnel can access the system
- ✓ **Data Protection and Privacy** – data needs to be protected (industrial property and for GDPR compliance)

5.2.3.7 (CPO1) COGNITIVE OPERATIONS SOLUTION



- ✓ **IIoT** – collection of data coming from various sensors and cognitive (CPS1-CPS5) sensors.
- ✓ **Data in Motion** – real time processing of operational data coming from OPC server. Alarms corresponding to potential faults to be visualized to inform the operator.
- ✓ **Situational Data** – management and aggregation of context data
- ✓ **Smart Data Space and Application** – cognitive control algorithm based on a predictive algorithm that visualizes potential future faults to the operator (according with real-time constraints this could be provided through edge nodes instead of the cloud).
- ✓ **Data at rest** – management of static data coming from internal and external sources
- ✓ **System Wrappers** – historical data coming from PLC and SCADA system can be acquired
- ✓ **Authorization and Authentication** – only authorized personnel can access the system
- ✓ **Identity Management** – only identified personnel can access the system
- ✓ **Data Protection and Privacy** – data needs to be protected (industrial property and for GDPR compliance)

5.2.3.8 (CPP1) COGNITIVE PLANNING SOLUTION



- ✓ **IIoT –** collection of data coming from various cognitive sensors (no real time data, but data captured during previous trials will be used)
- ✓ **Data in Motion –** real time processing of operational data (no real time data, but data captured during previous trials will be used)
- ✓ **Situational Data** – management and aggregation of context data
- ✓ **Smart Data Space and Application –** cognitive solution based on a process model developed in CPC1, with optimization of the scheduling of trials needed during process development.
- ✓ **Data at rest –** management of static data coming from internal and external sources
- ✓ **System Wrappers** – historical data coming from SCADA system can be acquired
- ✓ **Authorization and Authentication** – only authorized personnel can access the system
- ✓ **Identity Management** – only identified personnel can access the system
- ✓ **Data Protection and Privacy** – data needs to be protected (industrial property and for GDPR compliance)

5.2.4 Summary table

CAPRI RA	
Smart Data Spaces and Applications	
Data Access-Information Layer	CAS1, CAS2, CAC1, CAO1, CAP1, CSS1, CSS2, CSS3, CSS4, CSS5, CPS1, CPS2, CPS3, CPS4, CPS5, CPC1, CPO1, CPP1
Smart Data Management and Integration	
Cloud Computing-Industrial IoT and Analytics	CAS2, CAC1, CAO1, CAP1, CSS1, CSS2, CSS3, CSS4, CSS5, CPS1, CPS2, CPS3, CPS4, CPS5, CPC1, CPO1, CPP1 (*) most of them require running it in edge nodes
Situational Data	CAS1, CAS2, CAC1, CAO1, CSS1, CSS2, CSS3, CSS4, CSS5, CPS1, CPS2, CPS3, CPS4, CPS5, CPC1, CPO1, CPP1
Data in Motion (DiM)	CAS2, CAC1, CAO1, CAP1, CSS1, CSS2, CSS3, CSS4, CSS5, CPS1, CPS2, CPS3, CPS4, CPS5, CPC1, CPO1
Data at Rest (DaR)	CAS2, CAC1, CAO1, CAP1, CSS1, CSS2, CSS3, CSS4, CSS5, CPS1, CPS2, CPS3, CPS4, CPS5, CPC1, CPO1, CPP1
Smart field	
Devices	CAS2, CAC1, CAO1, CAP1, CSS1, CSS2, CSS3, CSS4, CSS5, CPS1, CPS2, CPS3, CPS4, CPS5, CPC1, CPO1
Sensors	CAS1, CAS2, CAC1, CAO1, CAP1, CSS1, CSS2, CSS3, CSS4, CSS5, CPS1, CPS2, CPS3, CPS4, CPS5, CPC1, CPO1
Actuators	CAS2, CAC1, CAO1, CSS5, CPS5, CPC1, CPO1
Machines	CAS2, CAC1, CAO1, CAP1, CSS1, CSS2, CSS3, CSS4, CSS5, CPS1, CPS2, CPS3, CPS4, CPS5, CPC1, CPO1
External Systems	
System Wrappers	CAS2, CAC1, CAO1, CSS1, CSS2, CSS3, CSS4, CSS5, CPS1, CPS2, CPS3, CPS4, CPS5, CPC1, CPO1, CPP1

Security	
Authorization and Authentication	all
Identity Management	all
Data Protection and Privacy	all
Data Sovereignty	
IDS App Store	TBD
IDS Clearing House	TBD
IDS Broker	TBD
IDES Connector	TBD

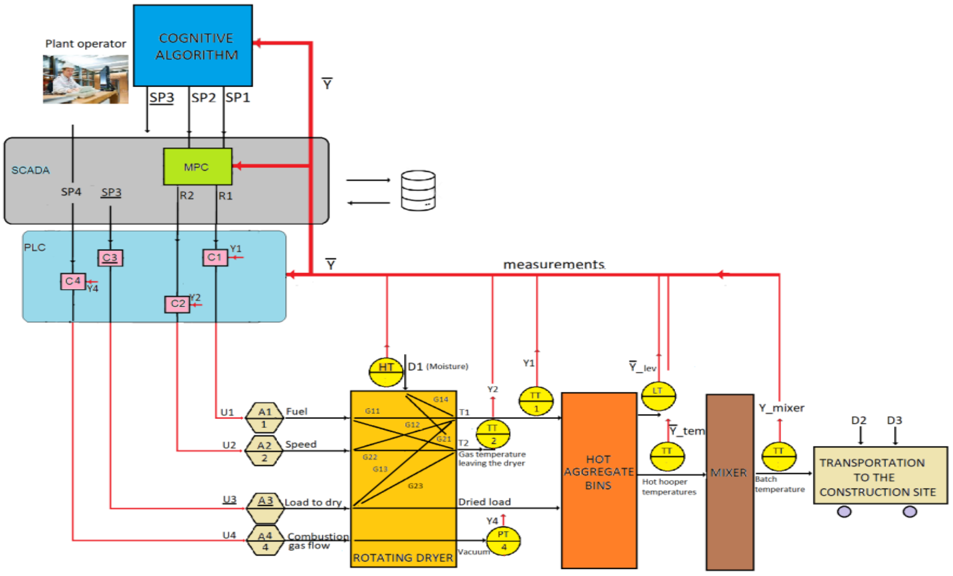
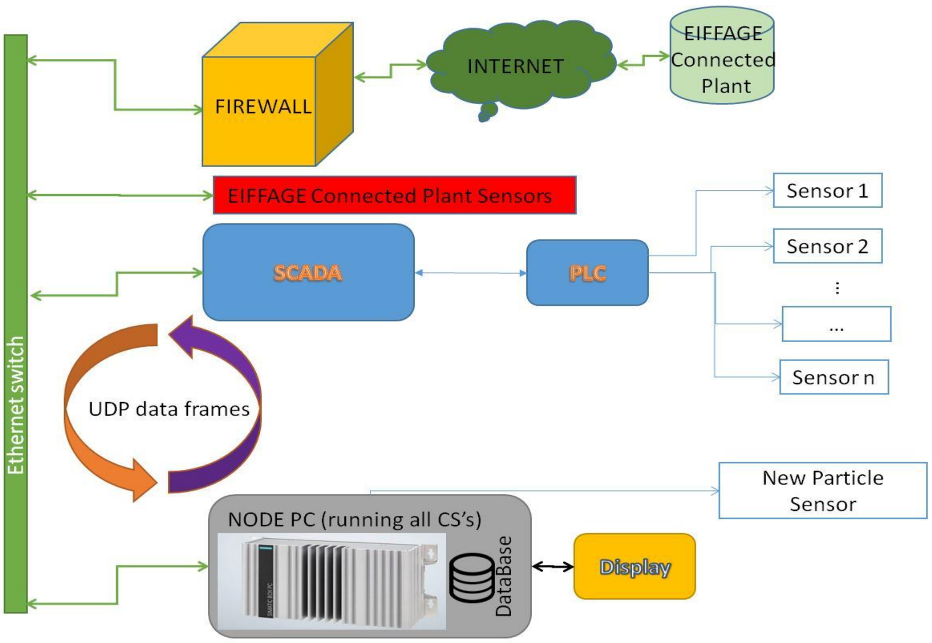
5.3 Mapping of the CAPRI Requirements

The present sub-chapter defines the mapping between the Reference Architecture described above and the requirements collected from the Case Studies and documented in D2.2, particularly the ones defining desiderata, assumptions and constraints that could have an impact from an architecture perspective.

With such a list of well-defined requirements, it is quite straightforward to check the soundness and completeness of the depicted Reference Architecture through the traceability matrixes listed below.

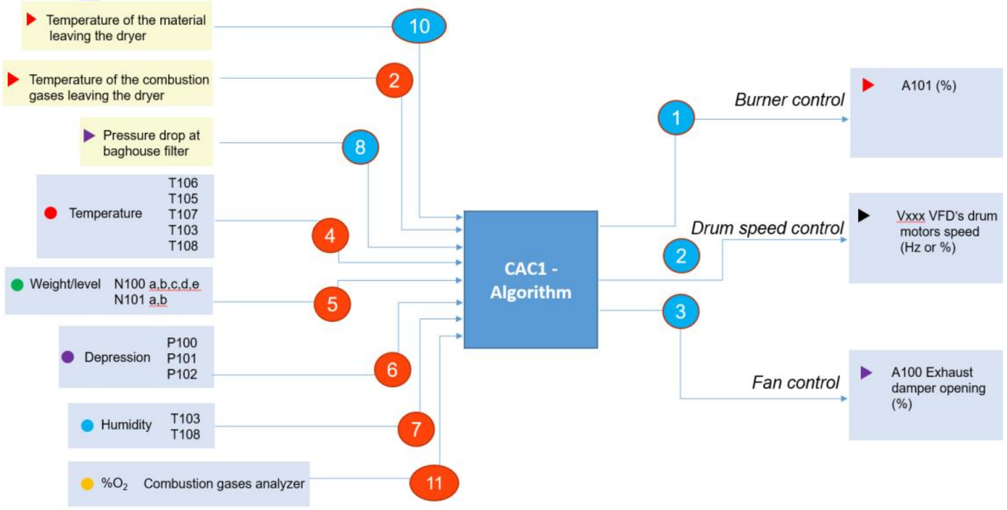
5.3.1 Asphalt

REQ CAC1 COGNITIVE CONTROL OF DRYER DRUM	
Smart Data Spaces and Applications	
Data Access-Information Layer	**CAC1 Cognitive Control Algorithm** based on an MPC control block running on a Node PC. **Functional Blocks:** Drying drum hybrid identified model MPC control block **Location:** Control cabin

<table>
<tr><td></td><td>
</td></tr>
<tr><td colspan="2">Smart Data Management and Integration</td></tr>
<tr><td>Cloud Computing-Industrial IoT and Analytics</td><td>EIFFAGE PLANT CONNECTED platform (MQTT Protocol: WAGO datalogger system sends data to the platform)</td></tr>
<tr><td>Situational Data</td><td>SQL database (TBD) running on a NODE PC
Location: Control cabin

</td></tr>
<tr><td>Data in Motion (DiM)</td><td>DiM → from SCADA system to NODE PC
UDP INTRAME proprietary dataframes in the LOCAL AREA NETWORK (Ethernet) (Location: control cabin):
Temperature sensors:
T106 Outlet temperature of the aggregates leaving the drying drum
T105 Temperature of the combustion gases leaving the drying drum</td></tr>
</table>

This project receives funding in the European Commission's Horizon 2020 Research Programme under Grant Agreement Number 870062

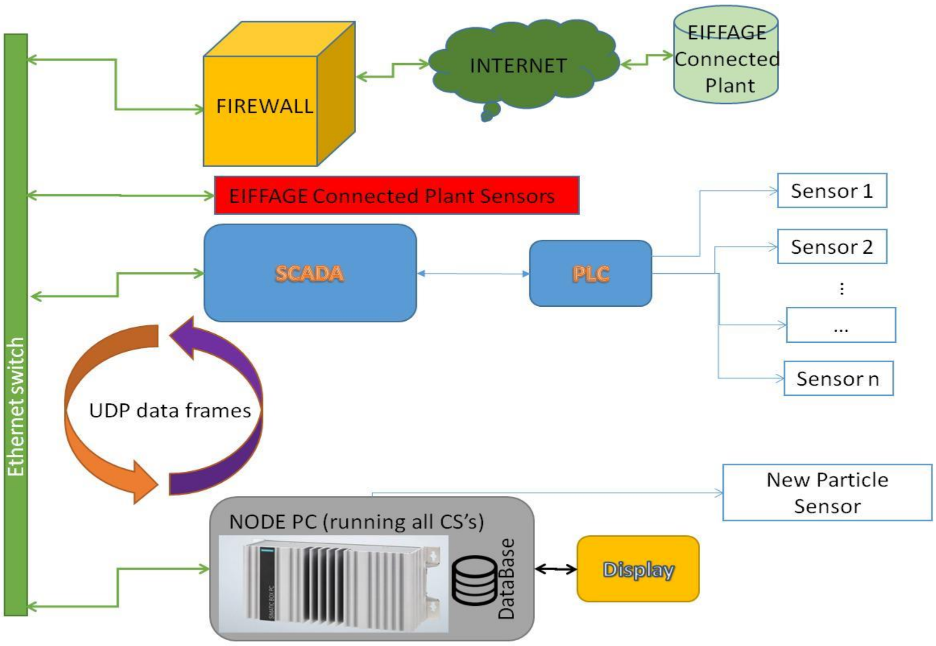
	T107 Temperature of the final asphalt mix in the mixer **Weight / level sensors:** N100 (a,b,c,d,e) Hot bins level (max.-min.) N101 (a,b) Filler silos level (max-min) Pressure sensors_ P100 Baghouse filter drop pressure P101 Drying drum to baghouse filter pipe drop pressure P102 Drying drum pressure **Humidity** T103 Humidity of the sands in the cold bins T108 Environmental Humidity **Gas composition** %O_2 combustion gases composition **DiM → From NODE PC to SCADA system** **UDP INTRAME proprietary dataframes in the LOCAL AREA NETWORK (Ethernet)** (Location: control cabin): **Burner control SETPOINT** A101 Burner power (%) **Drying drum speed SETPOINT** Vxxx VFDs drum motors speed (%) **Exhaust gases SETPOINT** A100 Exhaust damper opening (%) **System SETPOINTS** Temperature of the material leaving the dryer Temperature of the combustion gases leaving the dryer Drop pressure at the baghouse filter **DiM → from EIFFAGE CONNECTED PLANT system to NODE PC** **MQTT Protocol:** **WAGO datalogger system (proprietary EIFFAGE PLANT CONNECTED platform). It sends data to a cloud repository through a firewall.** Location: Control cabin **Sensors** **Temperature/Humidity** T103 Temperature and humidity of the sands in the cold bins T108 Environmental temperature and humidity

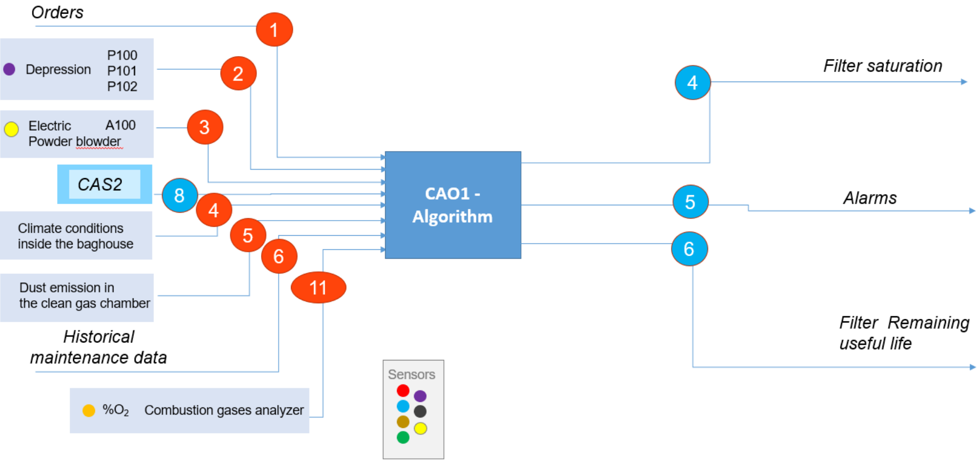
	
Data at Rest (DaR)	All previous data stored at an **SQL database** (TBD) running on a NODE PC **Location:** Control cabin
Smart field	
Devices	**UDP INTRAME proprietary dataframes in the LOCAL AREA NETWORK (Ethernet)** (Location: control cabin): **System SETPOINTS** Temperature of the material leaving the dryer Temperature of the combustion gases leaving the dryer Drop pressure at the baghouse filter
Sensors	**UDP INTRAME proprietary dataframes in the LOCAL AREA NETWORK (Ethernet)** (Location: control cabin): **Temperature sensors**: T106 Outlet temperature of the aggregates leaving the drying drum T105 Temperature of the combustion gases leaving the drying drum T107 Temperature of the final asphalt mix in the mixer. **Weight / level sensors:** N100 (a,b,c,d,e) Hot bins level (max.-min.) N101 (a,b) Filler silos level (max-min) Pressure sensors_ P100 Baghouse filter drop pressure P101 Drying drum to baghouse filter pipe drop pressure P102 Drying drum pressure **Gas composition** %O_2 combustion gases composition **MQTT Protocol:**

This project receives funding in the European Commission's Horizon 2020 Research Programme under Grant Agreement Number 870062

	WAGO datalogger system (proprietary EIFFAGE PLANT CONNECTED platform). It sends data to a cloud repository through a firewall. **Temperature sensors**: T103 Temperature of the sands in the cold bins T108 Environmental temperature **Humidity sensors** T103 Humidity of the sands in the cold bins T108 Environmental Humidity
Actuators	**UDP INTRAME proprietary dataframes in the LOCAL AREA NETWORK (Ethernet)** (Location: control cabin): **Burner control SETPOINT** A101 Burner power (%) **Drying drum speed SETPOINT** Vxxx VFDs drum motors speed (%) **Exhaust gases SETPOINT** A100 Exhaust damper opening (%
Machines	Rotary drying drum
External Systems	
System Wrappers	Historical data up to **TBD** days locally obtained through .CSV files from SCADA control computer and imported to NODE PC. Location: Control cabin Method: Manually
Security	
Authorization and Authentication	TBD
Identity Management	TBD
Data Protection and Privacy	TBD
Data Sovereignty	
IDS App Store	TBD
IDS Clearing House	TBD
IDS Broker	TBD
IDES Connector	TBD

REQ CAO1 PREDICTIVE MAINTENANCE OF BAGHOUSE BASED ON COGNITIVE SENSORS AND EXPERT KNOWLEDGE

Smart Data Spaces and Applications	
Data Access-Information Layer	**CAO1 Predictive Maintenance of Baghouse based on Cognitive Sensors** running on a Node PC. **Functional Blocks:** Different process measurements Connection with the Drum Cognitive Climatic conditions Wear of filters data Humidity. **Location:** Control cabin
Smart Data Management and Integration	
Cloud Computing-Industrial IoT and Analytics	**EIFFAGE PLANT CONNECTED platform (MQTT Protocol: WAGO datalogger system sends data to the platform)**
Situational Data	**SQL database** (TBD) running on a NODE PC **Location:** Control cabin 
Data in Motion (DiM)	**DiM → from SCADA system to NODE PC** **UDP INTRAME proprietary dataframes in the LOCAL AREA NETWORK (Ethernet)** (Location: control cabin): **Pressure sensors** P100 Baghouse filter drop pressure P101 Drying drum to baghouse filter pipe drop pressure

	P102 Drying drum pressure **Power consumption** A100: Powder blower electrical power consumption **Production orders** Production Orders (Recipes) **DiM → From NODE PC to SCADA system** **UDP INTRAME proprietary dataframes in the LOCAL AREA NETWORK (Ethernet)** (Location: control cabin): Filter saturation Alarms Filter lifetime **DiM → from EIFFAGE CONNECTED PLANT system to NODE PC** **MQTT Protocol:** **WAGO datalogger system (proprietary EIFFAGE PLANT CONNECTED platform). It sends data to a cloud repository through a firewall.** Location: Control cabin **Sensors** **Temperature / Humidity** T103 Temperature and humidity of the sands in the cold bins T108 Environmental temperature and humidity 
Data at Rest (DaR)	All previous data stored at an **SQL database** (TBD) running on a NODE PC **Database data in EXCEL spreadsheets** Location: Control cabin Maintenance history data Production orders **Location:** Control cabin

This project receives funding in the European Commission's Horizon 2020 Research Programme under Grant Agreement Number 870062

Smart field	
Devices	**UDP INTRAME proprietary dataframes in the LOCAL AREA NETWORK (Ethernet)** (Location: control cabin): **ALARMS / STATUS (to be sent to operator (SCADA) computer** Filter saturation Alarms Filter lifetime
Sensors	**UDP INTRAME proprietary dataframes in the LOCAL AREA NETWORK (Ethernet)** (Location: control cabin): **Pressure sensors** P100 Baghouse filter drop pressure P101 Drying drum to baghouse filter pipe drop pressure P102 Drying drum pressure **MQTT protocol** WAGO datalogger system (proprietary EIFFAGE PLANT CONNECTED platform) **Temperature** Temperature conditions within baghouse filter **Emissions** Dust emissions in the clean gas chamber
Actuators	**UDP INTRAME proprietary dataframes in the LOCAL AREA NETWORK (Ethernet)** (Location: control cabin): **Power consumption** A100: Powder blower electrical power consumption
Machines	Baghouse filter Exhaust fan/blower
External Systems	
System Wrappers	**Database data in EXCEL spreadsheets** Location: Control cabin Maintenance history data Production orders
Security	
Authorization and Authentication	TBD
Identity Management	TBD
Data Protection and Privacy	TBD
Data Sovereignty	

IDS App Store	TBD
IDS Clearing House	TBD
IDS Broker	TBD
IDES Connector	TBD

5.3.2 Steel

REQ CSS1 COGNITIVE STEEL TRACKING SENSOR	
Smart Data Spaces and Applications	
Data Access-Information Layer	Data access is realized via a webservice that abstracts away the low-level data acquisition protocols. **Functional Blocks:** - A laser engraves a QR-code in billets after the continuous caster **(requirement R1)** - A camera system to read the QR-code when a billet enters the reheating furnace for the hot rolling mill **(requirement R2)** - A laser engraces a QR-code in bars after the hot-rolling mill **(requirement R3)** - A camera system to read the QR-code of a bar that enters a finishing line **(requirement R4)**
Smart Data Management and Integration	
Cloud Computing-Industrial IoT and Analytics	Both the input data and the results of the cognitive sensor will be stored in the product digital twin and hence be amenable to analytics. The digital twin will provide an API to access this data for products under processing and historical ones.
Situational Data	
Data in Motion (DiM)	**Input data:** Per billet (from MES): - New billet marked (timestamp, id, heat id) **(requirements R1)** - Billet entering reheating furnace (timestamp, id, heat id) **(requirements R2)** Per bar (from MES): - New bar marked (timestamp, id, billet id, heat id) **(requirements R3)** - Bar entering finishing line (timestamp, id, billet id, heat id, finshing line id) **(requirements R4)**
Data at Rest (DaR)	Stored in an archive collection in a NoSQL database. **Location**: either on premise or in a public cloud
Smart field	
Devices	See Machines below

Sensors	See Data in Motion above
Actuators	None
Machines	• Continuous caster, equipped with laser engraving system for new billets • Hot rolling mill, equipped with camera scanner at the entrance of the reheating furnace for billet identification, and a laser engraving system for new bars. • Finishing lines, equipped with camera scanner for bar identification
External Systems	
System Wrappers	**MES**: data access
Security	
Authorization and Authentication	Not specified yet
Identity Management	Not specified yet
Data Protection and Privacy	Not specified yet
Data Sovereignty	
IDS App Store	Not specified yet
IDS Clearing House	Not specified yet
IDS Broker	Not specified yet
IDES Connector	Not specified yet

REQ CSS2 COGNITIVE SOLIDIFICATION SENSOR	
Smart Data Spaces and Applications	
Data Access-Information Layer	Data access is realized via a webservice that abstracts away the low-level data acquisition protocols.
Smart Data Management and Integration	
Cloud Computing-Industrial IoT and Analytics	Both the input data and the results of the cognitive sensor will be stored in the product digital twin and hence be amenable to analytics. The digital twin will provide an API to access this data for products under processing and historical ones.
Situational Data	
Data in Motion (DiM)	**Input data (requirements R1-R4, R6):** Per heat (from MES): • Steel analysis (chemical composition) • Stirring duration • Degassing duration • Deoxidation additions

	Periodic (interval in range 1s – 10s, from SCADA): - Casting velocity - Steel temperature at meniscus - Flow rate and temperature difference of cooling water in mould - Flow rates of spray water in different loops of secondary cooling Per billet (from SCADA): - Cutting length of billet **Output data (requirement R5):** - The temperature profile and solidification front per casting strand - All necessary temperature and upstream process information of the billets cut at the end of the caster
Data at Rest (DaR)	Stored in an archive collection in a NoSQL database. **Location**: either on premise or in a public cloud
Smart field	
Devices	See Machines below
Sensors	See Data in Motion above
Actuators	None
Machines	- Ladle furnace - Continuous caster
External Systems	
System Wrappers	**MES**: steelmaking data (heat data) **SCADA:** continuous caster machine data
Security	
Authorization and Authentication	Not specified yet
Identity Management	Not specified yet
Data Protection and Privacy	Not specified yet
Data Sovereignty	
IDS App Store	Not specified yet
IDS Clearing House	Not specified yet
IDS Broker	Not specified yet
IDES Connector	Not specified yet

REQ CSS3 COGNITIVE TEMPERATURE SENSOR	
Smart Data Spaces and Applications	
Data Access-Information Layer	Data access is realized via a webservice that abstracts away the low-level data acquisition protocols.

Smart Data Management and Integration	
Cloud Computing-Industrial IoT and Analytics	Both the input data and the results of the cognitive sensor will be stored in the product digital twin and hence be amenable to analytics. The digital twin will provide an API to access this data for products under processing and historical ones.
Situational Data	
Data in Motion (DiM)	**Input data:** Per billet (from Digital Twin): • Temperature profile after casting, from CSS2 **(requirement R1)** • Location history of billet (and bar), from CSS1 **(requirement R2)** Per billet (from MES): • Temperature scan of the billet after the reheating furnace **(requirement R3)** **Output data:** • The temperature evolution of the billet after casting and during rolling **(requirement R4)** • The temperature evolution of the bar after rolling and during the finishing **(requirement R5)**
Data at Rest (DaR)	Stored in an archive collection in a NoSQL database. **Location**: either on premise or in a public cloud
Smart field	
Devices	See Machines below
Sensors	See Data in Motion above
Actuators	None
Machines	• Continuous caster • Hot rolling mill • Finishing line
External Systems	
System Wrappers	**MES**: Access to temperature scan of the billet after leaving the reheating furnace (tbc) **SCADA:** TBC
Security	
Authorization and Authentication	Not specified yet
Identity Management	Not specified yet
Data Protection and Privacy	Not specified yet
Data Sovereignty	
IDS App Store	Not specified yet

IDS Clearing House	Not specified yet
IDS Broker	Not specified yet
IDES Connector	Not specified yet

REQ CSS4 COGNITIVE SCALE SENSOR	
Smart Data Spaces and Applications	
Data Access-Information Layer	Data access is realized via a webservice that abstracts away the low-level data acquisition protocols.
Smart Data Management and Integration	
Cloud Computing-Industrial IoT and Analytics	Both the input data and the results of the cognitive sensor will be stored in the product digital twin and hence be amenable to analytics. The digital twin will provide an API to access this data for products under processing and historical ones.
Situational Data	
Data in Motion (DiM)	**Input data:** Per billet (from Digital Twin): • Temperature evolution in the hot rolling mill, from CSS3 **(requirement R1)** Per billet (from MES (tbc)): • Force of work rolls, speed of material, width, thickness, flatness of products during/after the hot rolling **(requirement R2)** **Output data:** • An estimation of the total scale layer size **(requirement R3a)** • Optionally, in v2: an estimation of the scale layers’ chemical composition **(requirement R3b)**
Data at Rest (DaR)	Stored in an archive collection in a NoSQL database. **Location**: either on premise or in a public cloud
Smart field	
Devices	See Machines below
Sensors	See Data in Motion above
Actuators	None
Machines	• Hot rolling mill
External Systems	
System Wrappers	**Digital Twin:** Access to billet temperature evolution from CSS3 **MES**: Access to hot rolling mill process data
Security	
Authorization and Authentication	Not specified yet

Identity Management	Not specified yet
Data Protection and Privacy	Not specified yet
Data Sovereignty	
IDS App Store	Not specified yet
IDS Clearing House	Not specified yet
IDS Broker	Not specified yet
IDES Connector	Not specified yet

REQ CSS5 COGNITIVE RISK AND ANOMALIES SENSOR	
Smart Data Spaces and Applications	
Data Access-Information Layer	Data access is realized via a webservice that abstracts away the low-level data acquisition protocols. **Functional Blocks:** A data-driven model that estimates processing risks for steel billets and bars at multiple points in the processing chain, based on as many relevant input parameters as are available, including both raw process data and cognitive sensor results. **(requirement R3)** An anomaly detector trained and operating on the raw process data **(requirement R5)**
Smart Data Management and Integration	
Cloud Computing-Industrial IoT and Analytics	Both the input data and the results of the cognitive sensor will be stored in the product digital twin and hence be amenable to analytics. The digital twin will provide an API to access this data for products under processing and historical ones.
Situational Data	
Data in Motion (DiM)	**Input data:** Per billet/bar (from Digital Twin): • Cognitive sensor results from CSS1-CSS4 **(requirement R1)** Per bar (from MES (tbc)): • Quality labels for the finished product **(requirement R2)** **Output data:** • Processing risk per billet after casting and per bar after rolling **(requirement R4)** • Configurable anomaly alarms for anomalous events in the continuous casting, hot rolling and finishing stages **(requirement R6)**
Data at Rest (DaR)	Input data used for training (see DiM for required data points) made available via an archive collection of the digital twin. Likewise, results will be stored in an archive collection in the digital twin. **Location**: either on premise or in a public cloud

Smart field	
Devices	See Machines below
Sensors	See Data in Motion above
Actuators	None
Machines	• Continuous caster • Hot rolling mill • Finishing line
External Systems	
System Wrappers	**Digital Twin:** Access to cognitive sensor data and raw processing data
Security	
Authorization and Authentication	Not specified yet
Identity Management	Not specified yet
Data Protection and Privacy	Not specified yet
Data Sovereignty	
IDS App Store	Not specified yet
IDS Clearing House	Not specified yet
IDS Broker	Not specified yet
IDES Connector	Not specified yet

REQ DIGITAL TWINS	
Smart Data Spaces and Applications	
Data Access-Information Layer	Data access is realized via a webservice that abstracts away the low-level data acquisition protocols. **Functional Blocks:** • A product-centric view of the process data and cognitive sensor results. **(requirement R1)** • Visualizations **(requirement R2)** • Active replay simulation components and training for data-driven cognitive solutions (CSS4)
Smart Data Management and Integration	
Cloud Computing-Industrial IoT and Analytics	
Situational Data	

Data in Motion (DiM)	Currently active twins will be stored in a dedicated collection in order to guarantee fast access times and will be moved to an archive collection when finished.
Data at Rest (DaR)	See DiM.
Smart field	
Devices	See Machines below
Sensors	See Data in Motion above
Actuators	None
Machines	• Continuous caster • Hot rolling mill • Finishing line
External Systems	
System Wrappers	
Security	
Authorization and Authentication	Not specified yet
Identity Management	Not specified yet
Data Protection and Privacy	Not specified yet
Data Sovereignty	
IDS App Store	Not specified yet
IDS Clearing House	Not specified yet
IDS Broker	Not specified yet
IDES Connector	Not specified yet

5.3.3 Pharma

REQ CPS1 Blend uniformity	
Smart Data Spaces and Applications	
Data Access-Information Layer	Data access is realized via OPC server / client structure.
Smart Data Management and Integration	
Cloud Computing-Industrial IoT and Analytics	Process data will be stored in an SQL database system running SIPAT.
Situational Data	Ambient condition, granule properties
Data in Motion (DiM)	**Input data from OPC server running on Raman computer:** Raw spectra (~1800 points per spectrum, 1 spectrum per ~10 seconds) **Output data provided by OPC Server:**

	Computed API concentration
Data at Rest (DaR)	Stored at an **SQL database** (SIMATIC SIPAT) running on a NODE PC **Location:** Control cabin
Smart field	
Devices	Raman measurement system consisting of probe and measurement computer.
Sensors	PhAT Probe connected to the measurement computer
Actuators	None
Machines	After twin screw granulation
External Systems	
System Wrappers	OPC Server data needs to be accessed
Security	
Authorization and Authentication	Not specified yet
Identity Management	Not specified yet
Data Protection and Privacy	Not specified yet
Data Sovereignty	
IDS App Store	Not specified yet
IDS Clearing House	Not specified yet
IDS Broker	Not specified yet
IDES Connector	Not specified yet

REQ CPS2 Granule quality	
Smart Data Spaces and Applications	
Data Access-Information Layer	Data access is realized via OPC server / client structure.
Smart Data Management and Integration	
Cloud Computing-Industrial IoT and Analytics	Process data will be stored in an SQL database system running SIMATIC SIPAT.
Situational Data	Ambient condition
Data in Motion (DiM)	**Input data from OPC server running on Parsum computer:** Size raw data (~20 points per measurement, 1 measurement per second) **Output data provided by OPC Server:** Computed characteristic size descriptor

Data at Rest (DaR)	Stored at an **SQL database** (SIMATIC SIPAT) running on a NODE PC **Location:** Control cabin
Smart field	
Devices	Parsum measurement system consisting of probe and measurement computer.
Sensors	Parsum probe connected to the measurement computer
Actuators	None
Machines	After twin screw granulation
External Systems	
System Wrappers	OPC Server data needs to be accessed
Security	
Authorization and Authentication	Not specified yet
Identity Management	Not specified yet
Data Protection and Privacy	Not specified yet
Data Sovereignty	
IDS App Store	Not specified yet
IDS Clearing House	Not specified yet
IDS Broker	Not specified yet
IDES Connector	Not specified yet

REQ CPS3 Granule moisture	
Smart Data Spaces and Applications	
Data Access-Information Layer	Data access is realized via OPC server / client structure.
Smart Data Management and Integration	
Cloud Computing-Industrial IoT and Analytics	Process data will be stored in an SQL database system running SIMATIC SIPAT.
Situational Data	Ambient condition, granule properties
Data in Motion (DiM)	**Input data from OPC server running on ConsiGma computer:** Process data (100 data tags per measurement, 1 measurement per second) **Output data provided by OPC Server:** Granule moisture in the 6 dryer cells
Data at Rest (DaR)	Stored at an **SQL database** (SIMATIC SIPAT) running on a NODE PC **Location:** Control cabi

This project receives funding in the European Commission's Horizon 2020
Research Programme under Grant Agreement Number 870062

Smart field	
Devices	ConsiGma 25 manufacturing line
Sensors	Temperature, air flow, powder mass flow, liquid mass flow and humidity sensors
Actuators	None
Machines	Feeders, twin screw granulation and dryer
External Systems	
System Wrappers	OPC Server data needs to be accessed
Security	
Authorization and Authentication	Not specified yet
Identity Management	Not specified yet
Data Protection and Privacy	Not specified yet
Data Sovereignty	
IDS App Store	Not specified yet
IDS Clearing House	Not specified yet
IDS Broker	Not specified yet
IDES Connector	Not specified yet

REQ CPS4 Dissolution profile	
Smart Data Spaces and Applications	
Data Access-Information Layer	Data access is realized via OPC server / client structure.
Smart Data Management and Integration	
Cloud Computing-Industrial IoT and Analytics	Process data will be stored in an SQL database system running SIMATIC SIPAT.
Situational Data	Ambient condition, granule properties
Data in Motion (DiM)	**Input data from OPC server running on ConsiGma computer and on Parsum computer and on Raman computer:** Process data, size and concentration information (30 data tags per measurement, 1 measurement per second) **Output data provided by OPC Server:** Predicted dissolution profile
Data at Rest (DaR)	Stored at an **SQL database** (SIMATIC SIPAT) running on a NODE PC **Location:** Control cabin

Smart field	
Devices	ConsiGma 25 manufacturing line, Parsum measurement system, Raman measurement system, optionally automated tablet tester
Sensors	Temperature, powder mass flow, liquid mass flow, CPS1-CPS3 information
Actuators	None
Machines	ConsiGma 25 manufacturing line
External Systems	
System Wrappers	OPC Server data needs to be accessed, outputs of other cognitive solutions need to be accessed
Security	
Authorization and Authentication	Not specified yet
Identity Management	Not specified yet
Data Protection and Privacy	Not specified yet
Data Sovereignty	
IDS App Store	Not specified yet
IDS Clearing House	Not specified yet
IDS Broker	Not specified yet
IDES Connector	Not specified yet

REQ CPS5 Fault detection	
Smart Data Spaces and Applications	
Data Access-Information Layer	Data access is realized via OPC server / client structure.
Smart Data Management and Integration	
Cloud Computing-Industrial IoT and Analytics	Process data will be stored in an SQL database system running SIMATIC SIPAT.
Situational Data	Ambient condition, granule properties
Data in Motion (DiM)	**Input data from OPC server running on ConsiGma computer and on Parsum computer and on Raman computer:** Process data, size and concentration information (1000 data tags per measurement, 1 measurement per second) **Output data provided by OPC Server:** Indication of faults

Data at Rest (DaR)	Stored at an **SQL database** (SIMATIC SIPAT) running on a NODE PC **Location:** Control cabin
Smart field	
Devices	ConsiGma 25 manufacturing line, Parsum measurement system, Raman measurement system
Sensors	Temperature, powder mass flow, liquid mass flow, CPS1-CPS4 information
Actuators	None
Machines	ConsiGma 25 manufacturing line
External Systems	
System Wrappers	OPC Server data needs to be accessed, outputs of other cognitive solutions need to be accessed
Security	
Authorization and Authentication	Not specified yet
Identity Management	Not specified yet
Data Protection and Privacy	Not specified yet
Data Sovereignty	
IDS App Store	Not specified yet
IDS Clearing House	Not specified yet
IDS Broker	Not specified yet
IDES Connector	Not specified yet

REQ CPC1 Control concept	
Smart Data Spaces and Applications	
Data Access-Information Layer	Data access is realized via OPC server / client structure.
Smart Data Management and Integration	
Cloud Computing-Industrial IoT and Analytics	Process data will be stored in an SQL database system running SIMATIC SIPAT
Situational Data	Ambient condition, granule properties
Data in Motion (DiM)	**Input data from OPC server running on ConsiGma computer and on Parsum computer and on Raman computer:** Process data, size and concentration information (~ 10 data tags per measurement, 1 measurement per second for process control; ~30 data tags needed for material tracking) **Output data provided by OPC Server:**

	Liquid pump flow rate set point, dryer settings (air flow, air temperature, drying time) Discharge commands (product keys and tablets)
Data at Rest (DaR)	Stored at an **SQL database** (SIMATIC SIPAT) running on a NODE PC **Location:** Control cabi
Smart field	
Devices	ConsiGma 25 manufacturing line, Parsum measurement system, Raman measurement system
Sensors	Temperature, powder mass flow, liquid mass flow, CPS1-CPS4 information
Actuators	Liquid pump flow rate set point Dryer inlet air flow set point Dryer inlet air temperature set point Drying time Discharge command for product keys Discharge command for tablets
Machines	ConsiGma 25 manufacturing line
External Systems	
System Wrappers	OPC Server data needs to be accessed, outputs of other cognitive solutions need to be accessed
Security	
Authorization and Authentication	Not specified yet
Identity Management	Not specified yet
Data Protection and Privacy	Not specified yet
Data Sovereignty	
IDS App Store	Not specified yet
IDS Clearing House	Not specified yet
IDS Broker	Not specified yet
IDES Connector	Not specified yet

REQ CPO1 Operating solution	
Smart Data Spaces and Applications	
Data Access-Information Layer	Data access is realized via OPC server / client structure.
Smart Data Management and Integration	

This project receives funding in the European Commission's Horizon 2020 Research Programme under Grant Agreement Number 870062

Cloud Computing-Industrial IoT and Analytics	Process data will be stored in an SQL database system running SIMATIC SIPAT.
Situational Data	Ambient condition, granule properties
Data in Motion (DiM)	**Input data from OPC server running on ConsiGma computer and on Parsum computer and on Raman computer:** Process data, CPS5 output (~ 100 data tags per measurement, 1 measurement per second) **Output data provided by OPC Server and visualized:** Information on potential faults
Data at Rest (DaR)	Stored at an **SQL database** (SIMATIC SIPAT) running on a NODE PC **Location:** Control cabin
Smart field	
Devices	ConsiGma 25 manufacturing line, Parsum measurement system, Raman measurement system
Sensors	Temperature, powder mass flow, liquid mass flow, CPS1-CPS5 information
Actuators	None
Machines	ConsiGma 25 manufacturing line
External Systems	
System Wrappers	OPC Server data needs to be accessed, outputs of other cognitive solutions need to be accessed
Security	
Authorization and Authentication	Not specified yet
Identity Management	Not specified yet
Data Protection and Privacy	Not specified yet
Data Sovereignty	
IDS App Store	Not specified yet
IDS Clearing House	Not specified yet
IDS Broker	Not specified yet
IDES Connector	Not specified yet

REQ CPP1 Planning solution	
Smart Data Spaces and Applications	
Data Access-Information Layer	Data is obtained from csv files that have been exported from an SQL database

Smart Data Management and Integration	
Cloud Computing-Industrial IoT and Analytics	None
Situational Data	Ambient condition
Data in Motion (DiM)	None
Data at Rest (DaR)	Stored at an **SQL database** (SIMATIC SIPAT) running on a NODE PC; Will be exported to csv file and then read by CPP1 **Location:** Control cabin
Smart field	
Devices	None
Sensors	None
Actuators	None
Machines	ConsiGma 25 manufacturing line
External Systems	
System Wrappers	csv files need to be read
Security	
Authorization and Authentication	Not specified yet
Identity Management	Not specified yet
Data Protection and Privacy	Not specified yet
Data Sovereignty	
IDS App Store	Not specified yet
IDS Clearing House	Not specified yet
IDS Broker	Not specified yet
IDES Connector	Not specified yet

5.3.4 Summary Traceability Matrix

CAPRI RA			
Smart Data Spaces and Applications			
	Asphalt use case	Steel use case	Pharma use case
Data Access-Information Layer	CAS1-R1 CAS1-R2 CAS1-R4 CAS1-R5 CAS1-R6 CAS1-R8 CAS2-R1 CAS2-R2 CAS2-R5 CAS2-R6 CAS2-R7 CAS2-R8 CAS2-R9 CAC1-R1 CAC1-R2 CAC1-R3 CAO1-R1 CAO1-R2 CAO1-R3 CAP1-R1 CAP1-R2 CAP1-R3	CSS1-R1 CSS1-R2 CSS1-R3 CSS1-R4 CSS2-R1 CSS2-R2 CSS2-R3 CSS2-R4 CSS2-R5 CSS3-R1 CSS3-R3 CSS3-R4 CSS3-R5 CSS4-R2 CSS4-R3 CSS5-R3 CSS5-R4 CSS5-R5	CPS1-R1 CPS2-R2 CPS1-R3 CPS2-R3 CPS3-R1 CPS3-R2 CPS4-R1 CPS4-R2 CPS5-R1 CPC1-R1 CPC1-R2 CPC1-R3 CPC1-R4 CPO1-R1 CPO1-R2 CPP1-R1
Smart Data Management and Integration			
Cloud Computing-Industrial IoT and Analytics	CAS1-R1 CAS1-R2 CAS1-R6 CAS1-R8 CAS2-R1 CAS2-R2 CAS2-R5 CAS2-R6 CAS2-R7 CAS2-R8	CSS1-R1 CSS1-R2 CSS1-R3 CSS1-R4 CSS2-R1 CSS2-R2 CSS2-R3 CSS2-R4 CSS2-R5 CSS3-R1	CPS1-R1 CPS2-R2 CPS1-R3 CPS2-R3 CPS3-R1 CPS3-R2 CPS4-R1 CPS4-R2 CPS5-R1 CPC1-R1

	CAS2-R9 CAC1-R1 CAC1-R2 CAC1-R3 CAO1-R1 CAO1-R2 CAO1-R3 CAP1-R1 CAP1-R2 CAP1-R3	CSS3-R3 CSS3-R4 CSS3-R5 CSS4-R2 CSS4-R3 CSS5-R3 CSS5-R4 CSS5-R5	CPC1-R2 CPC1-R3 CPC1-R4 CPO1-R1 CPO1-R2 CPP1-R1
Situational Data	CAS1-R1 CAS1-R2 CAS1-R3 CAS2-R1 CAS2-R2	CSS1-R1 CSS1-R2 CSS1-R3 CSS1-R4	CPS1-R1 CPS2-R2 CPS1-R3
Data in Motion (DiM)	CAS1-R1 CAS1-R2 CAS1-R3 CAS1-R4 CAS1-R5 CAS2-R1 CAS2-R2 CAS2-R5 CAS2-R6 CAS2-R7 CAS2-R8 CAS2-R9 CAC1-R1 CAC1-R2 CAC1-R3 CAC1-R7 CAC1-R9 CAO1-R1 CAO1-R2 CAO1-R3 CAP1-R7	CSS1-R1 CSS1-R2 CSS1-R3 CSS1-R4 CSS2-R1 CSS2-R2 CSS2-R3 CSS2-R4 CSS2-R5 CSS3-R1 CSS3-R3 CSS3-R4 CSS3-R5 CSS4-R2 CSS4-R3	CPS1-R1 CPS2-R2 CPS1-R3 CPS2-R3 CPS3-R1 CPS3-R2 CPS4-R1 CPS4-R2 CPO1-R1 CPO1-R2
Data at Rest (DaR)	CAS2-R7	CSS1-R5	CPC1-R5

	CAO1-R6	CSS2-R6 CSS3-R6 CSS4-R4 CSS5-R3 CSS5-R4 CSS5-R5	CPP1-R1
Smart field			
Devices	CAS1-R1 CAS1-R2		
Sensors	CAS1-R1 CAS1-R2 CAS1-R5 CAS2-R1 CAS2-R2	CSS1-R1 CSS1-R2 CSS1-R3 CSS1-R4 CSS2-R1 CSS2-R2 CSS2-R3 CSS2-R4 CSS2-R5	CPS1-R1 CPS2-R2 CPS1-R3 CPS2-R3 CPS3-R1 CPS3-R2 CPS4-R1 CPS4-R2 CPS5-R1
Actuators	CAS1-R3 CAS2-R5		CPC1-R1 CPC1-R2 CPC1-R3 CPC1-R4
Machines			
External Systems			
System Wrappers	CAS2-R7 CAO1-R6 CAO1-R7		CPC1-R5 CPP1-R1
Security			
Authorization and Authentication			
Identity Management			
Data Protection and Privacy			
Data Sovereignty			
IDS App Store	TBD	TBD	TBD
IDS Clearing House	TBD	TBD	TBD
IDS Broker	TBD	TBD	TBD

IDES Connector	TBD	TBD	TBD

6 Conclusions

This deliverable starts to shape the CAPRI Reference Architecture taking into consideration the relevant requirements coming from the Case Studies. It offers a global view of the CAPRI RA providing information and complying to the emerging IIoT Reference Architectures and Models (IIRA, RAMI, FIWARE, BDVA, IDSA). It describes the functional layers that conform the Architecture and the components and technologies for the materialization of the cognitive automation reference implementation. The outcomes of WP2 will feed two parallel threads in WP3, the first, addressed in T3.5 and detailed in D3.1, oriented to provide a reference implementation of cognitive solutions for the process industry and the three use cases. The second thread is related to the technical developments and validation in industrial laboratory environment of cognitive technological solutions working independently, in WP3. Then, these technological solutions will be integrated within the cognitive automation platform in WP4, in real scenarios provided by the use cases, ending up with a tailoring activity to derive the reference blueprints for the use cases.

At the same time other architectures and technologies of other relevant SPIRE and Factories of the Future initiatives working on cognitive technologies and solutions have been analysed, and will be further investigated, with the aim to incorporate latest advancements in the cognitive solutions and lessons learnt from other research and innovation initiatives. This provides the basis, together with technical requirements identified by industrial experiments, for defining the final CAPRI Reference Architecture, covering also all sovereignty and security aspects.

The deliverable also provides detailed information on the proposed technologies to bring connectivity, interoperability, analytics and AI capabilities to the CAPRI platform, as well as the proposed Open Source components to cover them, mapping the full analytic stack in two parallel lanes of implementation based on FIWARE and APACHE components, to be further extended with the cognitive layer and to be deployed in edge and/or in the cloud, according with the requested performance constraints and response times.

In conclusion, the proposed CAPRI RA offers a good coverage for the whole requirements coming from the targeted sectors, as representative of the process industry, and paves the way for the successful implementation in coming work-packages of the projects (namely WP3, WP4 and WP5). The detailed analysis and mapping on functional and non-functional requirements coming from the three use cases and all proposed cognitive solutions makes full evidence of this. It is worth to notice that the Consortium is fully aware that some of the cognitive capabilities may rely on the emerging concept of the European Data Spaces and this will be for sure beneficial also in the process industry as such. To this end, the CAPRI RA has been designed to provide the elements to support tools and processes to assure Data Sovereignty and Data Economy, even if the current validation theaters are not currently focusing on cross-organization scenario, but some ongoing discussions (at Consortium level, but also in the other projects mentioned in Chapter 4.6 and 4.7, and in the wider SPIRE community) may have a positive impact in the future evolution and adoption of the CAPRI RA. This may generate new user needs, linked closely with data security, cloud interoperability, and artificial intelligence, so they will be object of refinement in the development of cognitive solutions steps (mainly in WP3 tasks) in the coming months.