A Reference Architecture for Data Quality in Smart Manufacturing

Francisco Fraile[1], Arcadio Garcia[2], Cinzia Rubattino[3], Sabrina Verardi[3] and Manfredi Giuseppe Pistone [3]

[1] Universitat Politècnica de València, Camino de Vera s/n, 46022, Valencia, Spain
[2] EXOS Solutions, S.L., Camino de Vera s/n, 46022, Valencia, Spain
[3] Engineering Ingegneria Informatica S.p.A., Piazzale dell'Agricoltura 24, 00144 Roma

Abstract

The i4Q Project aims to provide a complete set of solutions consisting of IoT-based Reliable Industrial Data Services (RIDS), the so-called 22 i4Q Solutions, able to manage the huge amount of industrial data coming from cheap cost-effective, smart, and small size interconnected factory devices for supporting manufacturing online monitoring and control. This papers presents the reference architecture used to guide the development of IoT-based Reliable Industrial Data Services for Manufacturing Quality Control. Based on industry standards and known best practices, the reference architecture adopts a three-tiered architectural model to represent the main system architectural components, and provides four different architectural viewpoints to address business, usage, functional, and implementation concerns.

Keywords

Data quality, industrial IoT, system architecture, reference models

1. Introduction

One of the challenges in implementing quality control processes and solutions is the development of a reference architecture technologies and on relevant sector specific standards. The Reference Architecture (RA) must not be designed from scratch, but must take into account the most recent releases of some relevant initiatives for digital industries, industrial IoT and edge computing - namely those from the Platform Industrie 4.0 initiative (RAMI 4.0) [1], the Industrial Internet Consortium (IIRA [2] and OpenFog RA [3]), the Big Data Value Association (BDVA RA) [4], Digital Factory Alliance Initiative (Digital Service RA) [5], the ISO Standard for the Internet of Things (ISO / IEC 30141) [6] – and contextualize them to the specific application domain. Besides, several research projects provide a wide set of background knowledge on the topic and represent solid backgrounds to stem upon.

This paper presents the second release of the i4Q reference Architecture, architecture for IoT-based Reliable Industrial Data Services (RIDS) [7, 8] ready to deal with large amounts of data in industrial environments, developed in the i4Q project [9].

2. Methodology

The i4Q Project aims to provide a complete set of solutions consisting of IoT-based Reliable Industrial Data Services (RIDS), the so-called 22 i4Q Solutions, able to manage the huge amount of

Proceedings of the Workshop of I-ESA'22, March 23–24, 2022, Valencia, Spain
EMAIL: ffraile@cigip.upv.es (F. Fraile); arcadio.garcia@exos-solutions.com (A. Garcia); cinzia.rubattino@eng.it (C. Rubattino); sabrina.verardi@eng.it (S. Verardi); manfredigiuseppe.pistone@eng.it (M.G Pistone)
ORCID: 0000-0003-0852-8953 (F. Fraile); 0000-0001-5370-0527 (C. Rubattino); 0000-0002-3563-8683 (S. Verardi)

industrial data coming from cheap cost-effective, smart, and small size interconnected factory devices for supporting manufacturing online monitoring and control.

The development of the i4Q reference architecture (i4Q RA) followed an architecture-centered, scenario-driven, iterative development process, based on the ISO/IEC/IEEE 42010 "Systems and software engineering – Architecture" [10], starting from an initial released that has been reviewed and refined in consecutive releases. The first release of the i4Q RA, was not designed from scratch, being strongly based on the most relevant outcomes of other previous research activities and releases of International Communities, compiled during the initial analysis of related standards and reference models. According to the analysis and the evaluation performed, it was mainly inspired by the Industrial Internet Reference Architecture (IIRA) architectural model, based on the three-tiers approach [11]. Also, according to the IIRA approach, the task of the i4Q RA definition was performed in parallel with those of its viewpoints analysis, namely the business viewpoint, usage viewpoint, functional viewpoint, and implementation viewpoint. The viewpoints offer a framework to think iteratively through the architectural issues that arose throughout the definition of the system architecture.

Since the first release, a strong connection between the RA and the 22 i4Q Solutions (17 tools and 5 guidelines) was defined. In order to have a coherent inclusive vision provided by the i4Q RA, a mapping activity of i4Q RIDS against it has been started since the beginning and feedback's collection from the solution providers were fundamental to refine the architecture and its building blocks, and to develop the different viewpoints.

This way, the architecture and the viewpoints are mutually correlated and continuous feedback from each other need to be exchanged following an iterative approach to guarantee the alignment between the different elements of the Reference Framework.

So, the analysis across the four key viewpoints served as input for the architecture, which at the same time influenced them. The combination of the results obtained from the different viewpoints was derived into detailed reference architecture. The results included: business, regulatory and stakeholders' key inputs (from the business viewpoint); an identification of tasks, roles or activities to be performed by the framework (from usage viewpoint); a decomposition into its Control, Operations, Information, Application and Business domains (from the functional viewpoint); an identification of associated flows and analysis and selection of the technologies required for its implementation (from implementation viewpoint).

Besides the architectural viewpoints, the i4Q RA also takes into account all the considerations coming from the use case scenarios, the requirements' collection and the specifications of the i4Q solutions. Special attention is devoted to the pilots' needs involving them in the activities of the viewpoints' definition and defining blueprints.

Another key input has been the definition of digital models and ontologies to be used in the i4Q Framework. The analysis of the reference models performed in the establishment of the first version of the reference architecture represented the basis for mapping activity to determine the best data models and ontologies for each architectural tier.

The definition of the use case scenarios and requirements analysis and functional specification paved the way for an in-depth analysis of the i4Q Solutions, performed in close collaboration with Solution Providers, to better understand how they fit into the RA and its different viewpoints. The architectural viewpoints, using the value-driven approach suggested in IIRA, were incrementally improved, enabling top-bottom (i.e., from the business perspective down to the technical implementation perspective) and bottom-up feedback loops organizing meetings to make sure that the different viewpoints are consistent. The operational needs characterizing the realization, deployment and use of each solution represented a valid input and helped to better identify the RA, refining its modules, components and functionalities.

The definition of the second version of the RA also took into account all the considerations formalized in the description of use case scenarios and the requirements' collection; special attention was paid to the needs emerging from the pilots, which have been mapped in detailed blueprints to verify the alignment between the specific components to be used for industrial problems and the RA building blocks.

It is therefore clear that the design of the i4Q RA has been an iterative and parallel process, a based on a mix of top-down and bottom-up approaches, in which the results provided by several activities have been gathered and included in a systematic project vision. All of the information gathered in this reference architecture lays the groundwork for future project steps.

3. Reference architecture

The i4Q Reference Architecture (i4Q RA) defines a conceptual framework aiming to be the canvas for the design, implementation and integration of the i4Q Solutions. Being strongly based on the most relevant outcomes of other previous Research and Innovation activities and releases of International Communities, the i4Q RA is aligned to the most common standard reference architectures in the manufacturing domain and is mainly inspired by IIRA architectural model [2].

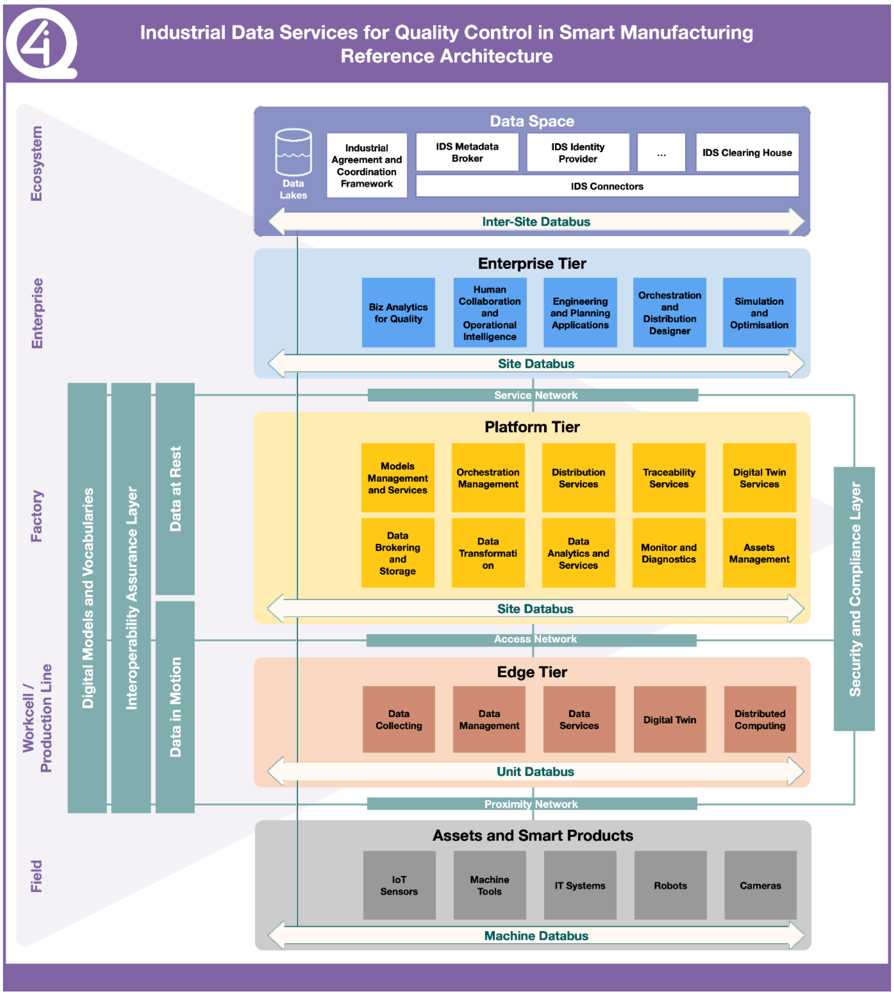
In particular, two approaches have been adopted:

- The IIRA three-layers architectural pattern as the main design driver.
- The IIRA layered databus pattern in order to make the solution data-driven.

The mapping between the layered databus pattern with the three-tier architecture pattern is as follows:

- The Machine Databus is mapped with the physical Assets and Smart Products.
- The Unit Databus is mapped with the Edge Tier.
- The Site Databus is mapped with the Platform and Enterprise Tiers.
- The Inter-Site Databus enables ecosystems and inter-factory communications.

The final version of i4Q RA fulfills the considerations coming from the viewpoints' developments, the pilots' needs and solution providers mapping activity.



. **Figure 1:** The i4Q Reference Architecture

The i4Q RA Security and Compliance Layer supports the communication between all tiers, providing a common security layer for accessing and retrieving data and physical assets (sensing and actuation).

Furthermore, the i4Q RA can support data (and services) description through the most significant meta- models (i.e., IIRA Characteristics of IIoT Information Models [12]), common ontologies and digital models. Finally, the Interoperability Assurance Layer guarantees interoperability inside the i4Q ecosystem considering the broad-spectrum of raw and elaborated data that can be processed by different solutions.

3.1. Edge tier

The Edge Tier aims to implement the sensor and control domain. It allows connecting physical assets and systems, collecting and processing data in the "proximity network". Furthermore, it supports Digital Twin (enabling advanced use, i.e., data filtering for generating a more precise virtual object, real-time analytics, etc.) and the Distributed Computing for sustaining distributed workloads and models. The Edge Tier, typically, process real-time data (Data in Motion).

The Edge Tier includes applications and services for describing and managing data, supports digital twins reducing the connectivity and latency issue, and the distributed computing devising a typical edge-cloud orchestration system.

Edge Tier consists of the following building blocks:

- Data Collecting: Solutions for data ingestion, collecting raw data from the facilities and store them to the data lake or make them available for the processing.
- Data Management: Solutions for data management, transformation (typically pre-processing), harmonization and loading.
- Data Services: Data Services to enable ingestion use and maintenance.
- Digital Twin: Digital Twin (Operational Services) to achieve a connected 3D production simulation with a digital twin for manufacturing.
- Distributed Computing: Fog Computing/Edge nodes that allow deploying and running AI workloads on the edge enable efficient analysis.

3.2. Platform tier

The Platform Tier contains applications and services for supporting the data flow (abstracted data) acting as a middleware between the Edge Tier and the Enterprise Tier (and between access and service network). It supports services for quality control (modelling, orchestration, traceability, etc.) and historical data processing (Data at Rest).

The Platform Tier includes subcomponents for integrating third-party systems and architectures, Digital Twin services for interacting with the edge-related application, and distribution services (i.e., resource management, data, and algorithms orchestration, etc.).

Platform Tier consists of the following building blocks:

- Models Management and Services: Management of Models and related Services enabling holistic vision of data across i4Q Infrastructure.
- Orchestration Management: Services for orchestration management enabling high resiliency and responsiveness workloads.
- Distribution Services: Distribution and deployment services enabling plug & play reconfiguration and installation of workloads.
- Traceability Services: Robust and Rapid Traceability Service (i.e., DLT) to provide an audit trail for all inserted data, guaranteeing immutability and finality.
- Digital Twin Services: Digital Twin (Management Services) services that enable industrial companies to virtual validation/visualization and productivity optimization using pre-existing and/or simulated data and data from different factory levels.

- Data Brokering and Storage: Services for data Brokering and storage management supporting a high degree of digitization in companies with most manufacturing devices acting as sensors or actuators and generating vast amounts of data.
- Data Transformation: Services for data transformation (typically post-processing).
- Data Analytics and Services: Services for data analytics on top of the data infrastructure with several incremental algorithms (i.e. operating on data streams with fast incremental updates) suitable for analytic processing of high-speed data streams.
- Monitor and Diagnostics: Services for near/real-time monitoring and diagnostics monitoring the health of workloads, predict problems and take corrective actions, predict failures and provide alerts.
- Assets Management: Services for assets management for managing and track facilities' assets.

3.3. Enterprise tier

The Enterprise Tier implements business-specific applications related to quality control, providing interfaces for end-users. It integrates engineering and management applications for supporting command generations, operational intelligence operations, big-data analytics and service orchestration.

The Enterprise Tier includes an "open" (not exhaustive) catalogue of use case/domain-specific applications. The usage of standard APIs and commons ontologies drives the definition of the common interfaces and facilitates the integration of (user and system) legacy and newer applications.

Enterprise Tier consists of the following building blocks:

- Biz Analytics for Quality: Applications built on top of underneath services (e.g: Models Management and Services, Data Analytics and Services) aiming to pursue excellence in quality focusing on zero-waste context.
- Human Collaboration and Operational Intelligence: CRM, CSM, DSS and ERP/MES and other services that provide smart alerting and quality diagnosis.
- Engineering and Planning Applications: CAD, CAE, CAM and AR/VR and everything related to the joint and integrated use of software systems for computer-aided design and computer-aided manufacturing.
- Orchestration and Distribution Designer: Engineering Environment for defining pipelines and distributions such as a scalable, easy to use, policy-based distribution mechanism to ease the task of distributing AI/ML models and other metadata to the edge.
- Simulation and Optimization: Simulation and Optimization applications for optimizing manufacturing processes.

3.4. Data space

The Data Space Tier extends and enhances interoperability with future manufacturing data spaces by the means of IDS Connectors, according to main principles defined by International Data Space Association (IDSA) [13].

The International Data Spaces (IDS) is a virtual data space that leverages existing standards, technologies and governance models well accepted in the data economy, to facilitate secure and standardize data exchange and data linkage in a trusted business ecosystem. It thus provides a foundation for facilitating innovative cross-company business processes and creating smart-service scenarios and guaranteeing data sovereignty for data owner [13]. i4Q Reference Architecture, thus, has been enriched with data sovereignty and sharing capabilities taking into account the needs of a constantly updated knowledge base to determine and present quality factors in manufacturing.

Data Space Tier consists of the following building blocks:

- Industrial Agreement and Coordination framework: A guidance to satisfy requirements in order to obtain a certification, including business, organizational and operational agreement.
- IDS Metadata Broker: A comprehensive connector that provides the necessary interfaces for communicating with any other International Data Spaces connector.

- IDS Identity Provider: An intermediary offering services to create, maintain, manage and validate identity information and for Participants in the International Data Spaces.
- IDS Clearing House: An intermediary providing clearing and settlements services for all financial and data exchange transactions within the International Data Spaces.
- IDS Connectors: A dedicated communication server for sending and receiving data in compliance with the general Connector specification.

4. Discussion

The design of the i4Q RA followed an iterative approach and the results provided by several activities have been gathered and included in a systematic project vision that is currently used to guide the development of the envisioned i4Q solutions.

The first version of i4Q RA has been taken as the basis, continuing the mapping activities of the i4Q Solutions against it and collecting feedback from solution providers. Another input for the architecture consisted in the analysis across the four key viewpoints (business, usage, functional and implementation), the architecture and the viewpoints are mutually correlated and continuous feedback from each other. The iterative approach followed for the implementation of the viewpoints was crucial to guarantee the alignment between the different elements of the Reference Framework.

Another key input is the definition of digital models and ontologies to be used in the i4Q Framework, of which analysis represents the basis for mapping activity to establish the best data models and ontologies for each architectural tier.

The definition of the second version of the i4Q RA has also taken into account activities carried out in the description of use case scenarios, the requirements' collection and the specifications of the i4Q RIDS. Special attention was devoted to the pilots' needs involving them in the activities of the viewpoints' definition and defining blueprints.

The second and final version of the reference architecture will serve as input for the implementation activities that will take place in i4Q project as well as sound basic for further European research projects.

5. References

[1] DIN SPEC91345:2016-04 - Reference Architecture Model Industrie 4.0 (RAMI4.0), 2016. URL: https://www.din.de/en/wdc-beuth:din21:250940128

[2] Industrial Internet Consortium, The Industrial Internet of Things - Volume G1: Reference Architecture, 2019. URL: https://www.iiconsortium.org/pdf/IIRA-v1.9.pdf

[3] OpenFog, OpenFog Reference Architecture for Fog Computing, 2017. URL: http://site.ieee.org/denver-com/files/2017/06/OpenFog_Reference_Architecture_2_09_17-FINAL-1.pdf

[4] BDVA, European BDVA Strategic Research and Innovation Agenda, 2017. URL: http://www.bdva.eu/sites/default/files/BDVA_SRIA_v4_Ed1.1.pdf

[5] Digital Factory Allience, Transforming Manufacturing Together. URL: https://digitalfactoryalliance.eu

[6] ISO, Internet of Things (IoT) - Reference Architecture, 2018. URL: https://www.iso.org/standard/50508.htm

[7] R. Poler, A. Karakostas, S. Vrochidis, A. Marguglio, S. Gálvez-Settier, P. Figueiras, ... C. Agostinho, An IoT-based Reliable Industrial Data Services for Manufacturing Quality Control, in: 2021 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), IEEE, New York, 2021, pp. 1-8. doi: 10.1109/ICE/ITMC52061.2021.9570203

[8] A. Karakostas, R. Poler, F. Fraile, S. Vrochidis, Industrial data services for quality control in smart manufacturing–the i4q framework, in: 2021 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4. 0&IoT), IEEE, New York, 2021, pp. 454-457. doi: 10.1109/MetroInd4.0IoT51437.2021.9488490.

[9] CORDIS, Industrial Data Services for Manufacturing Quality Control in Smart Manufacturing, 2021. URL: https://cordis.europa.eu/project/id/958205/en
[10] ISO, ISO/IEC/IEEE 42010:2011 Systems and software engineering - Architecture description, 2011. URL: https://www.iso.org/standard/50508.html
[11] S. W. Lin, B. Miller, J. Durand, R. Joshi, P. Didier, A. Chigani, ... G. Bleakley, Industrial Internet Reference Architecture, 2015. URL: https://hub.iiconsortium.org/iira#:~:text=Industrial%20Internet%20Reference%20Architecture%20(IIRA)&text=The%20IIRA%20is%20a%20standards,guide%20technology%20and%20standard%20development.
[12] IIC, Characteristics of IIoT Information Models, 2020. URL: https://www.iiconsortium.org/pdf/Characteristics-of-IIoT-Information-Models.pdf
[13] IDSA, Reference Architecture Model, 2019. URL: https://internationaldataspaces.org/