D4.12 – i4Q Analytics Dashboard V2

WP4 – BUILD: Manufacturing Data Analytics for Manufacturing Quality Assurance

Document Information

GRANT AGREEMENT NUMBER	958205	ACRONYM	i4Q
FULL TITLE	Industrial Data Services for Quality Control in Smart Manufacturing		
START DATE	01-01-2021	DURATION	36 months
PROJECT URL	https://www.i4q-project.eu/		
DELIVERABLE	D4.12 – i4Q Analytics Dashboard V2		
WORK PACKAGE	WP4 – BUILD: Manufacturing Data Analytics for Manufacturing Quality Assurance		
DATE OF DELIVERY	CONTRACTUAL	31-Dec-2022	ACTUAL	30-Dec-2022
NATURE	Report	DISSEMINATION LEVEL	Public
LEAD BENEFICIARY	UNINOVA		
RESPONSIBLE AUTHOR	UNINOVA		
CONTRIBUTIONS FROM	1-CERTH, 2-ENG, 5-KBZ, 7-IKER,		
TARGET AUDIENCE	1) i4Q Project partners; 2) industrial community; 3) other H2020 funded projects; 4) scientific community		
DELIVERABLE CONTEXT/ DEPENDENCIES	This document is D4.12. It presents the second version of i4Q Analytics Dashboard. First version was presented in D4.4 - i4Q Analytics Dashboard.		
EXTERNAL ANNEXES/ SUPPORTING DOCUMENTS	None		
READING NOTES	None		
ABSTRACT	Deliverable "D4.12 - i4Q Analytics Dashboard V2" is a technical specification document, that covers the technical and development aspects of the i4Q Analytics dashboard solution (I4Q^{AD}). It describes in detail the role, the functionalities, and the conceptual architecture of I4Q^{AD}.		

Document History

VERSION	ISSUE DATE	STAGE	DESCRIPTION	CONTRIBUTOR
0.1	7-Nov-2022	ToC	Table of Contents	UNINOVA
0.2	22-Nov-2022	Draft	First Draft of the Deliverable to be sent to the consortium	UNINOVA, KBZ, IKER, CERTH, ENG
0.3	02-Dec-2022	Internal Review	Internal Review	IVLAB, KBZ
0.4	09-Dec-2022	Final draft	Address comments from internal reviewers	UNINOVA
1.0	30-Dec-2022	Final doc	Quality check and issue of final document	CERTH

Disclaimer

Copyright message

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS/ACRONYMS

AD	Analytics Dashboard
CD	Continuous Delivery
CI	Continuous Integration
DA	Data Analytics
ORM	Object-relational Mapper
PDF	Portable Document Format
RA	Reference Architecture
SQL	Structured Query Language
UI	User Interface

Executive summary

This document presents an executive explanation of the **i4Q Analytics Dashboard** (I4Q^{AD}) Solution providing the general description, the technical specifications and the implementation status. It is an updated deliverable with respect to its previous version D4.4-i4Q Analytics Dashboard. The deliverable D4.12 is the Source Code of the i4Q^{AD} Solution that is in a private repository of Gitlab: https://gitlab.com/i4q.

The documentation associated to the i4Q^{AD} Solution is deployed on the website http://i4q.upv.es. This website contains the information of all the i4Q Solutions developed in the project "Industrial Data Services for Quality Control in Smart Manufacturing" (i4Q). The direct link to the i4Q^{AD} Solution documentation is **http://i4q.upv.es/12_i4Q_AD/index.html**

Such documentation is structured according to:

- General description
- Features
- Images
- Authors
- Licensing
- Pricing
- Installation requirements
- Installation Instructions
- Technical specifications of the solution
- User manual

Document structure

Section 1: Contains a general description of the **i4Q** **Analytics Dashboard**, providing an overview and the list of features. It is addressed to final users of the i4Q Solution.

Section 2: Contains the technical specifications of the **i4Q** **Analytics Dashboard**, providing an overview and its architecture diagram. It is addressed to software developers.

Section 3: Details the implementation status of the **i4Q** **Analytics Dashboard**, explaining the current status, and summarizing the implementation history.

Section 4: Provides the conclusions.

APPENDIX I: Provides the PDF version of the **i4Q** **Analytics Dashboard** web documentation, which can be accessed online at: **http://i4q.upv.es/12_i4Q_AD/index.html**

1. General Description

1.1 Overview

i4Q^{AD} is a reporting interface that allows monitoring industrial data with fully flexible visualisation drill-down charts and a flexible dashboard to provide meaningful analytics to users in a real-time basis using incremental algorithms. This i4Q solution aims to provide a graphical interface where it will be possible to visualise any kind of Data and/or other types of information provided by other solutions.

This solution has the main function of providing visual analytics tools and methods to the i4Q project. The i4Q Analytics Dashboard can be used via a Web Application or through the provision of a deployment bundle that can be deployed on-premises or on the cloud, and which is based on Apache Superset, an open-source software cloud-native application for data exploration and data visualisation able to handle data at petabyte scale.

1.2 Features

- Providing visual analytics tools and methods
- Support real-time stream data visualisation
- Feasibility of receiving input data in various formats and connecting to most of the SQL based data sources through SQLAlchemy
- Ability to manage users authentication and data asset permissions
- Customised visualisation dashboard based on i4Q look and feel
- Alerting system to handle the signal and alerting generation for threshold reaching

2. Technical Specifications

2.1 Overview

The i4Q **Analytics Dashboard** (i4Q^{AD}) aims to provide a graphical representation of extensive input data and information provided by other solutions. This solution relies on Apache Superset (https://superset.apache.org/) as the main visualisation technology since it covers the end user's desired visualisations as well as the technical requirements.

The solution is mapped to the Monitor and Diagnosis sub-component of the Platform Tier and to the Business Analytics for Quality of the Enterprise Tier, since it provides facilities to visualise and analyse multiple data sources at the same time including input data and the information provided by the other solutions, as well as monitor the quality of the processes for the end-users.

2.2 Architecture Diagram

i4Q^{AD} solution includes processes and services which are mapped to the Enterprise and Platform Tiers of the i4Q Reference Architecture. Sub-components of the i4Q Architectural Framework presented in D2.7. Considering that, this solution is found at various levels of the architecture:

- **Enterprise Tier:** The i4Q^{AD} mapping to the "Biz Analytics for Quality" subcomponent. It provides a visualisation tool capable of delivering analytics to manage the monitoring of processes and quality indicators in real-time, this helps manufacturers gain insight in specific operations, to aid in the decision-making process.
- **Platform Tier:** The i4Q^{AD} mapping to the "Monitor and diagnostics" subcomponent provides an easy way for a user to monitor and optimise production efficiency. By using multiple visualisation techniques, this tool will provide more flexibility in terms of the types of visualisations that can be presented, making it more adaptable and allowing for other components like the Data Analytics and services or the digital twin services to utilise their visualisation features with ease.
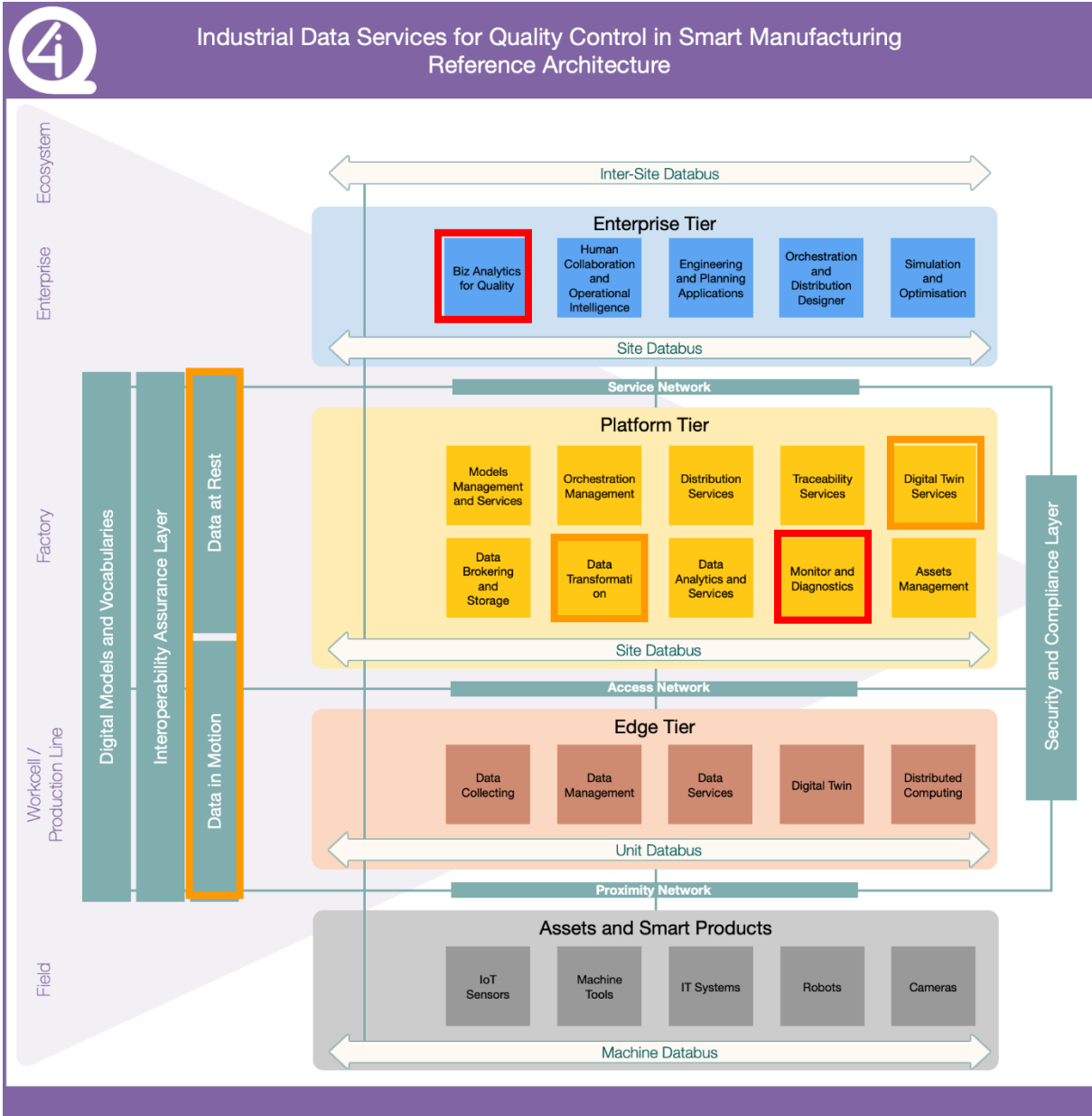
Figure 1. i4Q RA mapping with i4Q^AD

3. Implementation Status

3.1 Current implementation

The i4Q **Analytics Dashboard (**i4Q^{AD}) is a visualisation dashboard, its current implementation is based on Apache Superset. The dashboard offers multiple features such as visualising data in different graphs and charts, various kinds of transformation and filtering, exploring the data with running SQL queries, etc. The i4Q^{AD} uses Python ORM (SQLAlchemy) which supports connection with most SQL databases.

The current implementation of i4Q^{AD} contains:

- Customised version of Apache Superset dashboard with generalised i4Q theme.
- CI/CD that builds the Analytics Dashboard with the latest changes.
- Docker containers for each service that are easily distributable across platforms that support docker.
- Ability to connect data provider service with multiple data streaming brokers at the same time.

The current development is mainly based on two important additional features to Apache superset: 1) User interface customisation and 2) real-time data streaming; which are explained in more details in the following.

3.1.1 Analytics Dashboard UI customisation

i4Q^{AD} UI Customisations in Superset UI are done according to the i4Q look and feel. The following list includes some of the of justifications that are done for this regard:

- UI colors matching with general i4Q theme
- All logos and artworks are replaced with i4Q^{AD} logos and artworks including loading screen animation, header logo and lazy loading animation
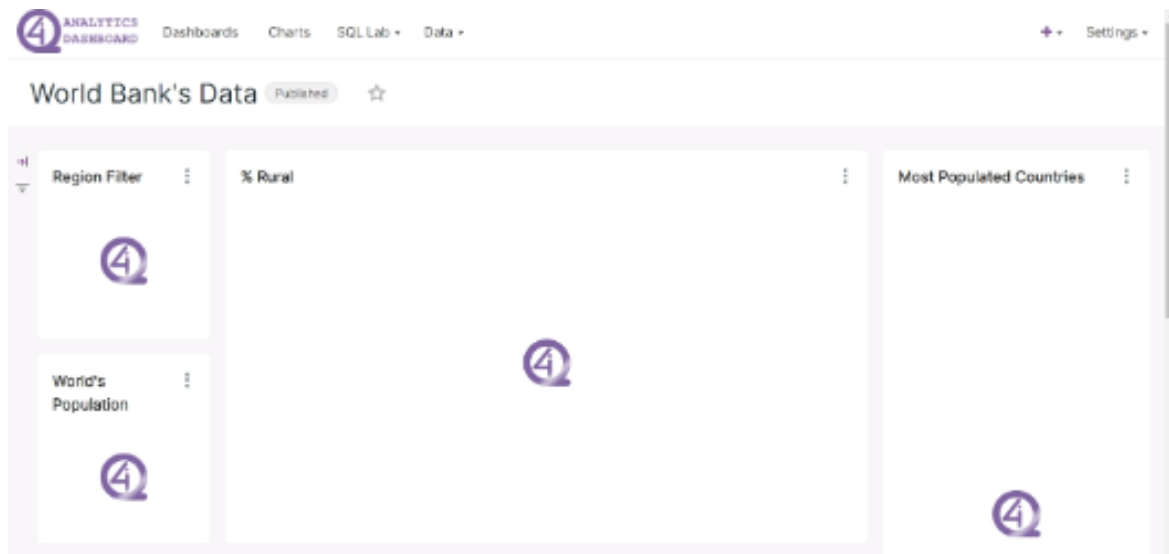- Configuration for adjusting specific features needed for i4Q^{AD}

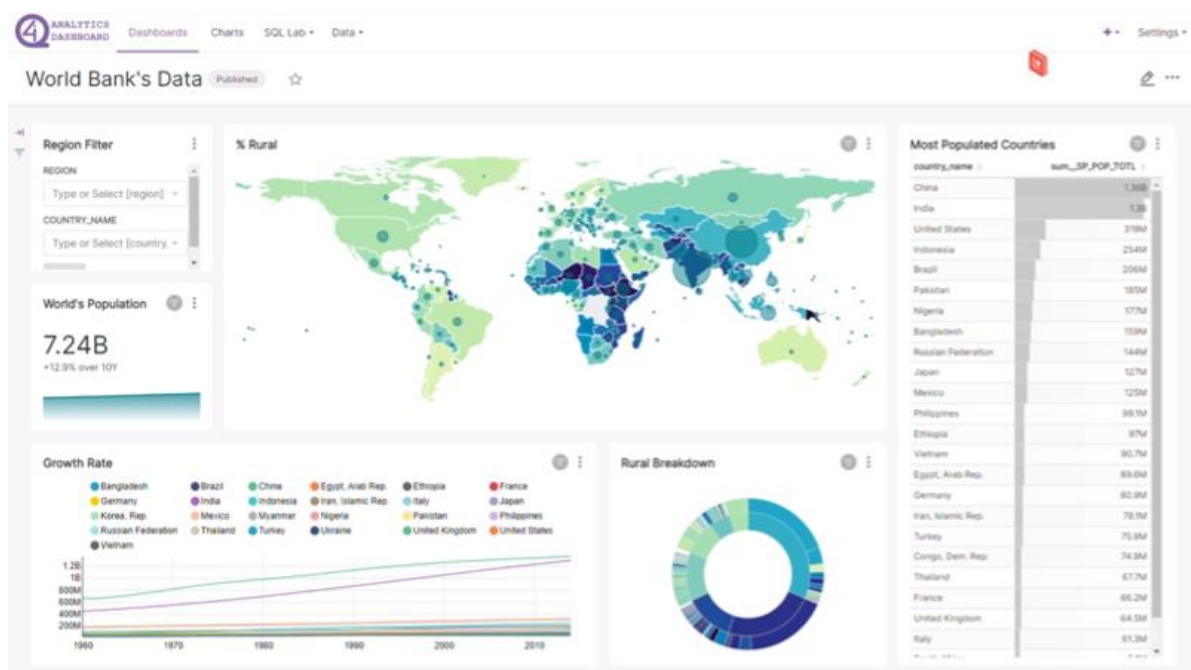**Figure 2.** i4Q Analytics Dashboard loading charts

Figure 3. i4Q Analytics Dashboard with multiple sample charts

3.1.2 Data Streaming Feature

Although Apache superset is a powerful visualisation tool it doesn't support data streaming. In order to deal with real-time data visualisation, the data streaming support is provided to i4Q Analytics Dashboard by using the Druid service (https://druid.apache.org/) as a middleware that processes data streams from Apache Kafka broker (https://kafka.apache.org/) and makes them available for real-time visualisation.

Druid and Kafka connection has been performed which allows the production of events through a Kafka producer where they are buffered in streams before they are used by Druid as a consumer in real-time. Druid provides Kafka indexing service that reads directly from Kafka. Streaming

ingestion uses an ongoing process called a supervisor that reads from the data stream to ingest data into Druid. Druid can replay events if the ingestion pipeline ever fails in some way. Kafka and Druid are used together to make the data available for visualisation in the real-time streaming analytics platform.

Druid has several process types, briefly described below:

- Coordinator processes manage data availability on the cluster.
- Overload processes control the assignment of data ingestion workloads.
- Broker processes handle queries from external clients.
- Router processes are optional; they route requests to Brokers, Coordinators, and Overloads.
- Historical processes store queryable data.
- MiddleManager processes ingest data.

The data stream from Kafka is parsed, transformed, and tuned in one-time configuration performed using Druid web console to make it available for being ingested by the analytics dashboard.
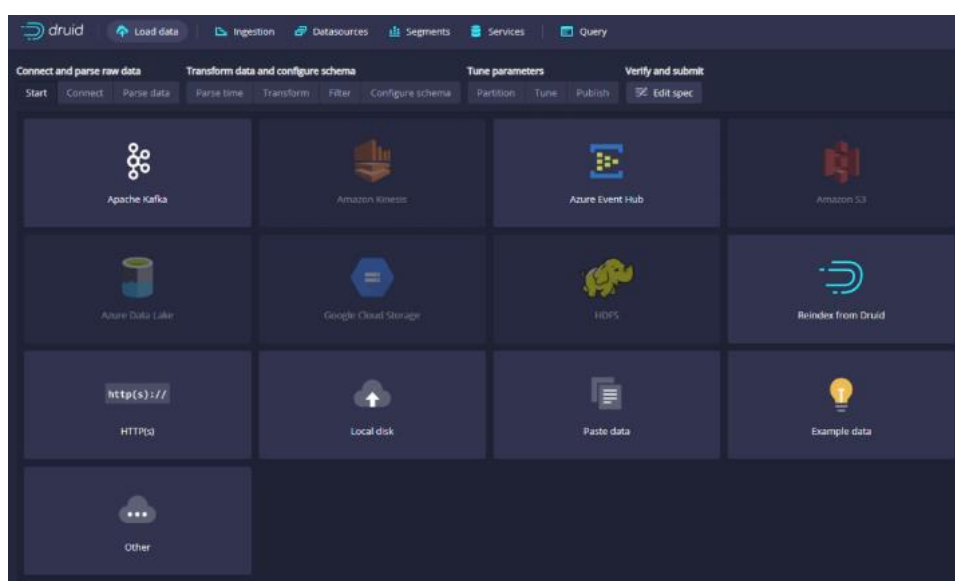

Figure 4. Druid web console

The i4Q^{AD} solution is containerised using Docker containers which allows easy distribution of solutions across multiple platforms. The docker services include 10 containers: Kafka, Superset, Zookeeper, PostgreSQL, and six Druid containers.

- **Compose file**: The docker-compose will create a container for Kafka, Superset, each Druid service, as well as ZooKeeper and a PostgreSQL container as the metadata store. A shared storage volume called druid_shared is also created which acts as persistent storage to keep and share segments and logs among Druid services. The volume is mounted at path opt/shared in file system of the container.
- **Configuration**: Configuration of the i4Q^{AD} docker containers is done via environment variables set within the containers. For Druid, the docker Compose passes the values from the Druid environment file into the container. The variables may additionally specify paths to the standard Druid configuration files which must be available within the container.

3.1.3 Solution features analysed and mapping with user requirements

A set of features has already been developed for i4Q^{AD}, based on the set of user requirements referring to i4Q^{AD} (see Deliverable 1.9) and in line with the functional viewpoints (presented in Deliverable 2.6). Similar requirements have been assigned into common categories of tasks based on an extensive technical study conducted on user requirements, introduced to ensure the generalisation abilities of the i4Q^{AD} solution.

In the following, we explain in more detail which solution's features address and cover each one of the requirements defined in D1.9 for the i4Q^{AD}:

- PC1r3 "Extract new features from collected data" is covered by the development of the Kafka-Druid-Superset pipeline which makes the solution capable of applying transformation and filtering on the input data in the pipeline.
- PC1r6.4 "Capability of displaying all information generated by the system related to failing components and protect equipment" is covered by Kafka-Druid-Superset pipeline that allows receiving the data and the other types of information provided by other solutions for the visualisation purpose.
- PC3r4.1 "Dashboard and Data Visualisation" is covered by using Superset as the base service that provides the feasibility of creating various types of charts and graphs.
- PC3r4.2 "Dashboard Usability" is covered by using Superset as the base service that provides a user-friendly interface.
- PC4r5.4.1 "Extract valuable info from existing data" is covered by the solution's capabilities to query the input data using SQL queries which allows extracting valuable information for the visualisation.
- PC5r2.2 "Visualisations of realtime/proactive analysis on the raw matter's composition and granulometry" is covered by the data streaming feature that enables the solution to connect the message broker and receive data streams in real time.
- PC6r3.1 "Create an interface to visualise and import information and visualise and send signals": i4Q^{AD} solution has the capability of importing information from different data sources and visualising in different charts and tables. The signal and alert generation will be activated in future developments.

3.2 History

Version	Release date	New features
V0.1	21/03/2022	Initial build
V0.2	22/03/2022	Added i4Q theme customisations
V0.3	03/04/2022	Added docker services
V0.4	04/04/2022	Added Druid and Kafka support
V0.5	06/04/2022	Revised docker configuration
V0.6	07/04/2022	Updated installation guidelines
V0.7	08/04/2022	Added Gitlab registry image
V1.0	07/05/2022	M18 Solution Release
V1.1	25/09/2022	Added User Authentication Role for i4Q pilots
V1.2	01/11/2022	Updated installation guidelines
V1.3	15/11/2022	Alerting system activation
V1.4	30/11/2022	Update Gitlab registry image
V2.0	15/12/2022	M24 Solution Release

Table 1. i4Q^{AD} Version history

4. Conclusions

Deliverable "D4.12 - i4Q Analytics Dashboard V2" is a technical specification document, that covers the technical and development aspects of the i4Q Analytics dashboard solution (i4Q^{AD}). It describes in detail the role, the functionalities, and the conceptual architecture of i4Q^{AD}. It presents a study detailing the main features of the solution to clarify the key functionalities and objectives of the i4Q^{AD} solution, describing its architecture diagram with respect to i4Q Reference Architecture.

Furthermore, the current implementation status of i4Q^{AD} has been explained in detail, presenting the significant progress of this overall development. This document presents the analysis of the pilot's requirements and the definition of the inputs and outputs of the solution.

Appendix I

The PDF version of the **i4Q** **Analytics Dashboard** web documentation can be accessed online at:
http://i4q.upv.es/12_i4Q_AD/index.html