Semantic Context Broker Tool

30.12.2022

Deliverable Number: D2.2

Authors:
Jordi Ricard Onrubia Palacios & Aitor Corchero

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

Deliverable No.	D2.2
Deliverable nature	OTHER (Other Services)
Work Package (WP) and Task	WP2/Task 2.2
Dissemination level	PU
Number of pages	44
Keywords	Semantic context brokers, Open data spaces, ontologies
Authors	Jordi Ricard Onrubia (EUT), Aitor Corchero (EUT)
Contributors	Andrea Marinoni (UiiT), Lydia Vambakeridou-Lyroudia (KWR), Dionysis Nikopoulos (NTUA), Sandra Banusch (KWB), Laia Romero (LOB), Nicolo Franceschetti (MGIS), Paris Chrysos (MANTIS)
Contractual submission date	30.11.2022
Actual submission date	30.12.2022

[1] PU = Public
CO = Confidential, only for members of the consortium (including the Commission Services)
R = Report
ORDP = Open Research Data Pilot

Technical References

Project acronym	IMPETUS
Project full title	Dynamic Information Management Approach for the Implementation of Climate Resilient Adaptation Packages in European Regions
Call	H2020-LC-GD-2020-2
Grant number	101037084
Project website	http://climate-impetus.eu/
Coordinator	EUT

Document history

V	Date	Beneficiary	Author
V1.0	06.10.2022	EUT	Jordi Ricard Onrubia Palacios
V1.2	08.11.2022	EUT	Jordi Ricard Onrubia Palacios
V1.2	25.11.2022	EUT	Aitor Corchero Rodriguez
V2.0	13.12.2022	EUT	Aitor Corchero Rodriguez
V2.1	15.12.2022	EUT	Jordi Ricard Onrubia Palacios & Aitor Corchero Rodriguez
V2.2	29.12.2022	NTUA (Quality Assurance)	Dionysios Nikolopoulos
FINAL	30.12.2022	EUT (Finalisation)	Hannah Arpke

Disclaimer

Abbreviations

Abbreviation / Acronyms	Description
(A)MGA	(Annotated) Model Grant Agreement
CA	Consortium Agreement
CFS	Certificate of Financial Statement
EAB	External Advisory Board
EC	European Commission
EU	European Union
FP	Framework Programme
GA	Grant Agreement
PSB	Project Steering Board
PMT	Project Management Team
PC	Project Consortium
WP	Work Package
WPL	Work Package Leader
WEFE	Water-Energy-Food-Ecosystems
RKB	Resilience Knowledge Boosters
JSON-LD	JavaScript Object Notation – Linked Data
API	Application Programming Interface
NGSI-LD	Next Generation Service Interfaces – Linked Data
CRUD	Create, Read, Update, Delete
DB	Database
FAIR	Findability, Accessibility, Interoperability, and Reusability
URI	Uniform Resource Identifier

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

Table of contents

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

List of figures

List of tables

Table of listings

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

Executive Summary

This deliverable is aimed at describing the elaborated **initial version of the semantic context broker** to (i) integrate climatic regional, national, and transnational data to tackle further analysis and predictions; (ii) enable data interoperability between public repositories, local/regional data and, specific key community systems data; (iii) support data understanding and management according to FAIR principles using meta-data descriptions.

According to the elaboration of the **Resilient knowledge booster digital dimension** (WP2) and the corresponding architecture described in the D2.1 ((UiT), 2022) the semantic context broker present a key role in the data management strategy. This tool **support on the elaboration of adaptation pathways and interventions in the case studies through the data understanding. Also, the tool will sustain the implementation of artificial intelligence tools to permit policy makers to understand climatic patterns and interrelation between climatic events and regional, national and transboundary information**.

Hence, the semantic context broker is one of the main pieces (see Figure 1) to support the relevant stakeholders (defined in D2.1) to combine different sources of information in such a way supports the analysis of climatic policies and goals in a wide-range of regions considering different time-scales.

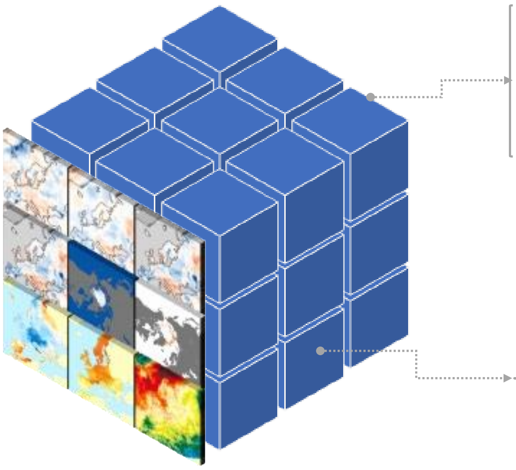

Figure 1. Semantic Context Broker vision

The highlighted part of the deliverable focus on the **description of the initial prototype of the semantic context broker that establishes the technological bases for data integration from case-studies combined with the integration of data from public and private repositories**. This initial prototype is available under the following URL:

http://84.88.176.103:10002

All the aforementioned work is related to the activities of the same WP, receiving input from T2.1 Design strategy of the digital dimension of Resilience Knowledge Boosters, WP3 Exposure and Vulnerability Assessment, specifically from T3.1 Generation of weather and climate data and T3.6 Strategic Resilience and Multi-Hazard Management tool for identifying dynamic adaptation pathways, WP4 Deployment of Solutions at demo sites and WP5 IMPETUS Adaptation Pathways and Innovation Packages. Furthermore, it will also provide output to support tasks from the same WP, task T3.5 Analyse and assess costs, benefits and risks related to interventions from WP3, all tasks from WP4, T5.1 Adaptation Pathways and T5.3 Evaluation of regional Innovation Packages from WP5 and finally T6.2 Provision of briefings and guidelines for decision-makers in industry and finance and T6.3 Exploitation of IMPETUS solutions from WP6 Boosting project impact.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

1 Introduction

IMPETUS project aims to **develop and validate a coherent multi-scale, multi-level, cross sectoral climate change adaptation framework to accelerate the transition towards a climate-neutral and sustainable economy**. To achieve this goal, IMPETUS will implement the so-called **"Resilient Knowledge Boosters" (RKB) (see** Figure 2**)** as a space where **all relevant stakeholders gather to co-create, demonstrate, monitor, and assess the climate adaptation pathways for sustainable adaptation and resilience**. To support the stakeholder's decision-making and subsequently, the generation of the adaptation strategies and interventions in the different regionalities, the RKBs will provide an online tool (virtual dimension) where relevant climate-change related data, knowledge and experiences will be integrated in a virtual and digital place to promote data/knowledge exchange between stakeholders and regions during the co-creation process.

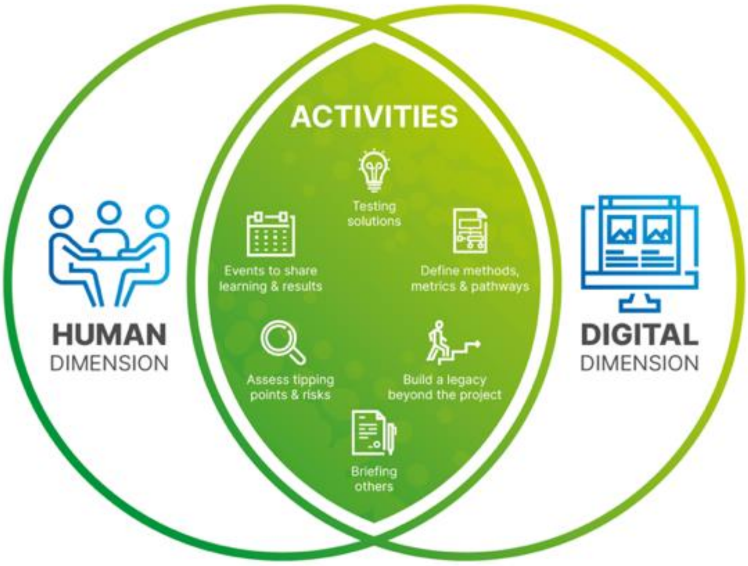

Figure 2. General Concept for the Resilient Knowledge Booster

In detail, the RKBs will be developed as **a main asset for the different IMPETUS case-studies** and support the different communities of the case-studies to tackle holistic, join and collaborative decisions. Based on this, the RKBs implementation will require of constant involvement of the case-studies and their related communities to:

1. **Understand their policy goals and climatic objectives** to focus the design of the online tool based on their main interests and challenges.
2. **Collect most representative data from the regionalities** that will be combined with public repositories.
3. **Integration of information coming from regional digital tools** such as modelling tools, simulation tools, decision-making tools, early warning systems and digital twins that permit the fusion of regional information and analysis of private and small data interrelationships between different key community systems.
4. **Provide feedback about the different visualizations and user interactions** to maximise the chances of the adoption of the RKBs in the regions but also, to expand the concept to other similar biogeographical regions (transferability and replicability).

Therefore, a robust stakeholder engagement methodology (WP1) is necessary to maintain the active involvement of the case studies' stakeholders in the RKBs' elaboration. For that, WP1 will work closely with technical WPs (WP2, WP3 and WP5) to support the integration of the different digital tools such as: (i) the semantic context broker; (ii) the visual and data exploration tool of the RKBs; (iii) the artificial intelligence tools; (iv) climatic scenarios; (v) selection of relevant KPIs for each of the regions; (vi)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

simulation and modelling tools; and (vii) the Climate change adaptation strategies and innovation packages for the regions.

The main IMPETUS objectives and challenges, rely on the RKBs framework. The relevant task regarding RKBs' conceptualization started on Month 6 (M6) with the identification of the requirements and the initial design of the RKBs. Results from this initial task were covered under D2.1, entitled "*Design strategy of the digital dimension of Resilience Knowledge Boosters*" ((UiT), 2022) that was delivered in M12. Several parallel tasks, related with the RKBs' continuing implementation, followed up. These tasks correspond with the implementation and elaboration of the semantic context broker (present deliverable) that comprises the work performed between M8 and M14 in the development of digital tools and the corresponding REST services to integrate and expose data coming from the IMPETUS case-studies. Also, the development of the RKBs comprises the generation of weather data and predictions aimed at providing climatic projections and forecasts to the case-studies (Task 3.1) and the integration of the 'hot-spot' tool facilitates the data fusion of main public repositories at European level in relation to climate and environment data (ERA5, CIMP6, COPERNICUS and IIASA repositories) (Task 3.3).

This afore mentioned information, will be integrated inside the **semantic context broker** to connect EU public information with public/private information coming from the Demo Sites. These digital tools are mainly devoted to **integrate information from multiple sources coming from multiple geographical scales (communities, regions, national and transboundary datasets) as well as from multiple-temporal scales (short, medium and long term)**. Indeed, the semantic context broker ensures **data interoperability, management according to the FAIR principles, and establishes a data framework to consume data using existing data standards**.

Thus, this online tool is responsible for managing the different sub-layers integrating **the physical layer** including all regional, national, and transnational datasets for each of the bioclimatic regions that takes part in the project. The tool manages the **data layer** in charge of persisting the collected data for short term view (consumption) but also, for the long-term analysis including data curation/wrangling. Also, the tools provide needed information for the **visualization layer** to permit an understanding of the climatic patterns at different scales, understanding of climatic events and the causal interrelation of anthropogenic factors. At the **application layer**, the tool establishes easy communication (via Open API) to enable the integration of third-party applications. Hence, the context broker is FIWARE-compatible; it complies both with the southbound API (i.e., collection of data from data repositories and external tools) and the northbound API (i.e., exposition of curated data to other systems).

The main novelty in the use of context brokers is their capacity to **harmonise the data capture and exposition through the use of common data exchange models based on the standard ETSI-GS-CIM 009 (well known as a NGSI-LD)** (ETSI, 2019). Thus, the context broker enables the description and registering of data and data repositories using data models (termed smart data models) that represents the information collected. This data-model-driven approach to define the digital tool provide several benefits such as: (i) data integration, harmonization and interoperability at cross-domain level; (ii) data visualization and exploration using standard data models structure and common data exchange languages (based on JSON-LD); (iii) Open APIs based on standards for both data integration from external repositories and also, data exposition and interlink with potential third parties applications; (iv) enabling the reuse of data integration tools (called IoT Agents) and then, minimising the elaboration of data-intensive applications.

The elaboration of the semantic context broker is performed under WP2,"*Digital and knowledge dimension of the Resilience Knowledge Boosters*". Specifically, the work is performed under Task 2.2, "*Open data space for enabling knowledge sharing*". This task is devoted to build a digital tool (context broker) to integrate information from multiple sources and enable federated data sharing across different disciplines. Accordingly with this main objective, the present deliverable is focused on presenting a general description of the architecture to position the context broker tool. Moreover, this deliverable will present also the mechanistic and detailed modules that takes part of the context broker. Indeed, the present deliverable will specify a REST API (based on the NGSI-LD standard) that will allow the management of the data, enabling the reception of the data from the different data sources to be stored and later shared among the different layers and services of the RKB architecture.

The present document gives an in-depth description of the semantic context broker and the integration capabilities it offers. the rest of the document is structured in the following manner:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

- **Chapter 2**: Review of the available context brokers as well as the most relevant open data spaces initiatives present at European Level.
- **Chapter 3**: Presents a brief introduction to the platform architecture and the WPs interrelation with this architecture.
- **Chapter 4**: In-depth description of the IMPETUS context broker and the integration capabilities it offers.
- **Chapter 5**: Demonstrates the deployment of a proof-of-concept context- broker prototype.
- **Chapter 6**: Outlines the conclusions and future work.

2 Common Initiatives towards the implementation of environmental and climatic open data space

This part of the document is focused on providing an overview of the main initiatives and trends in Europe in relation to reference data architectures and the so-called new wave of "Open Data Spaces". Both architectures are mainly devoted to the free flow of data across EU countries as described in the EU Artificial Intelligence Act (European Commission, Directorate-General for Communications Networks, Content and Technology, 2018) published on 2018 and its revision published on 2021 in a document entitled "Proposal for a regulation laying down harmonised rules on artificial intelligence" (European Commission, Directorate-General for Communications Networks, Content and Technology, 2018). This new trend and policy actions will establish the basis for the future artificial intelligence systems in Europe tackling relevant aspects such as:

1- **Data integration, harmonization and interoperability** between different countries and regions.
2- **Trustworthy Artificial Intelligence** as a mechanism to generate robust, reliable, and confident AI algorithms.
3- **Ethical and non-discriminatory AI** to ensure the data and the subsequent artificial intelligence algorithms that exploit such data are FAIR and do not have any undesirable bias in the information.
4- **Data Governance and Sovereignty** to ensure data privacy and security that permit to make the data available and sharable considering the different country laws ecosystem and the ownership of such data.

Considering these established principles in the European Union, this chapter comprises the overview of Open Data Space initiatives (Section 2.1) as a reference architecture to share data. Complementing this vision and as a part of such reference architectures, an overview of available context brokers are explained under the Section 2.2.

2.1 Open Data Spaces and Reference architectures to share climatic and environmental information

Cross-domain data availability in Europe and Worldwide at has been increased, notably due to the novel data infrastructures, distributed data architectures and the operationalization of monitoring and sensing tools. The huge amount of data available, combined with the complexity and heterogeneity of data sources, data sharing flows and concerns about data ownership privacy and security, have imposed the need to invest in data-driven innovations and data interoperability approaches. In this direction, the European Union (EU), defined the policy vision and framework for enabling data drive-society in the future. Thus, the policy framework in relation to data and data-driven innovations comprises in 2020 the publication of the European Strategy for data (European Commission, Directorate-General for Communications Networks, Content and Technology, 2020) whose main purpose is to improve the use of data and the subsequent interoperability across domains and countries. Following up this strategy, the Data Governance Act (European Commission, Directorate-General for Communications Networks, Content and Technology, 2020) was published in 2020 describing the corresponding processes and structures to facilitate voluntary data sharing by different societal pillars (private sector, public sector, and citizens). The Data Act (European Commission, Directorate-General for Communications Networks, Content and Technology, 2020) established the rules and conditions for data governance, fairness and sovereignty and finally, the Open Data Directive (European Parliament, Council of the European Union, 2019) exposes the technical requirements and datasets that public sector bodies are required to publish in a machine-readable format. Under the combination of these directives and policy-based documents, the European Data Space concept appears to remove the barriers and silos caused by the collection of information from different processes at domain-specific level (agriculture, water, energy, food, climate, etc). The concept of European Data Space also intends to remove the siloes caused by the lack of data sharing at cross-domain level. As explained in (Edward Curry S. S., 2022)], the Big Data Value Association published the main principles and future directions of the data spaces as represented in the Figure 3.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

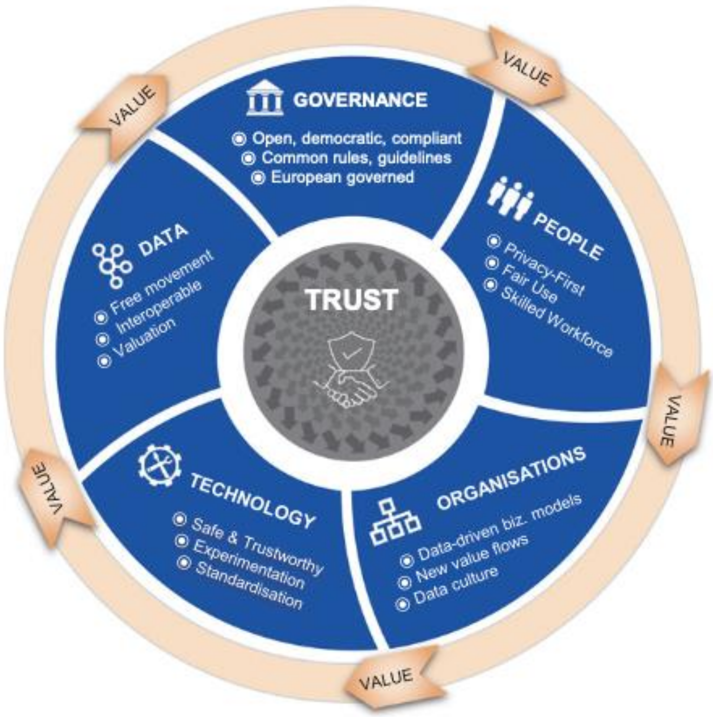

Figure 3. European Data Spaces principles and concept at general scale (Edward Curry S. S., 2022)

Considering these principles and concepts, several initiatives appeared in Europe in relation to the Open Data space market. These initiatives correspond to the International Data Spaces Association (IDSA), GAIA-X, SIMPL, Open DEI, Data Space Business alliance (DSB), FIWARE and iSHARE approaches. Due to the heterogeneity in dataspaces that appeared at European level, the rest of the section will describe each of the initiatives, provide a comparison, and list the challenges in their implementation.

2.1.1 Overview of Open Data Spaces Initiatives in Europe

2.1.1.1 International Data Spaces Association (IDSA)

The International Data Spaces Association (IDSA) is an initiative made up of more than 130 organizations from more than 20 countries, which was launched by the Fraunhofer Institute in 2015. IDSA has defined a reference architecture and a formal standard that will be used to create and operate virtual data spaces. The IDS initiative aims to create secure and trusted data spaces, where companies of any size and from any industry can sovereignly manage their data assets.

The IDS architecture, called IDS-RAM, is based on a commonly accepted data governance model that facilitates the secure and reliable exchange, and easy linking of data within enterprise ecosystems. The IDS-RAM architecture (Figure 4) corresponds to a high-level, peer-to-peer architecture model where data can be shared across interested third parties. Thus, this architecture guarantees that data providers can make data available for exchange and sharing, maintaining their digital sovereignty. This therefore forms the basis for developing and offering intelligent services and establishing innovative business processes.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

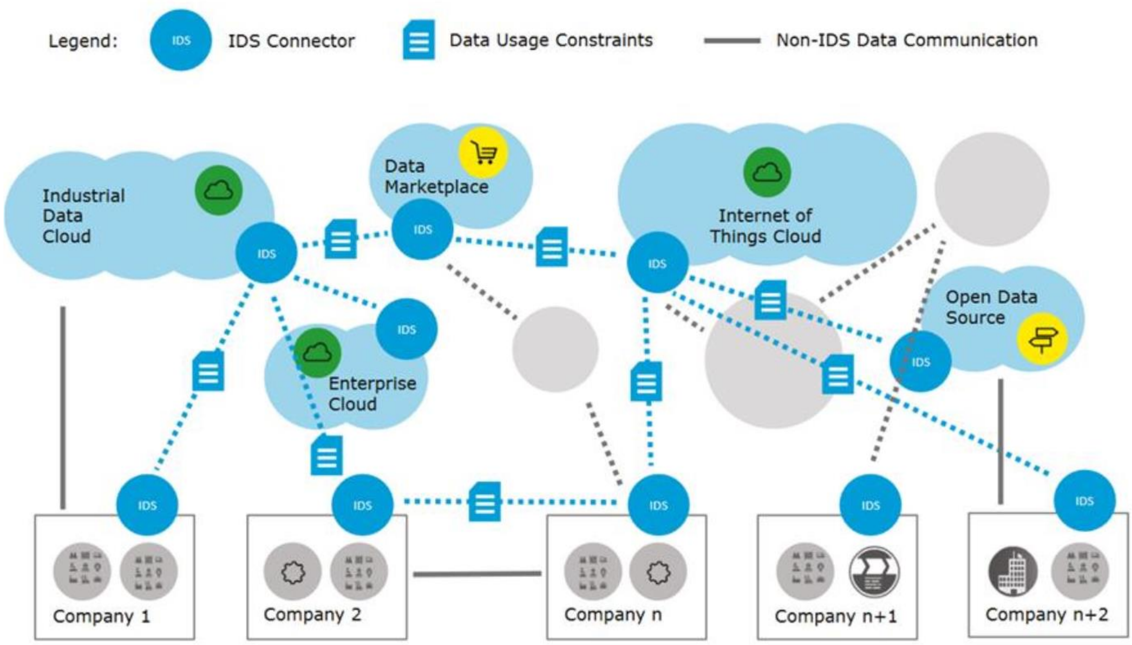

Figure 4. IDSA-RAM Architectural model

Complementing the IDS-RAM, the IDS initiative has published a series of catalogues with examples of implementation of data spaces that use its architecture and the components defined in it, many of which come from national and European R&D projects.

2.1.1.2 GAIA-X

The GAIA-X initiative emerged from the Franco-German public-private ecosystem in 2019, although it was not until February 2021 that it took shape as a pan-European (non-profit) association based in Brussels. The initiative is articulated around the two axes of the European Data Strategy: the data economy and cloud services (more generically, on the flexibility and ease provided by the cloud as-a-service model).

The goal of GAIA-X is the development of an open software layer of control and governance, and the implementation of a common set of policies and rules that will apply to any existing cloud/edge technology stack for transparency, sovereignty, and interoperability between data and services. Its deployment could be made from any cloud service provider that implements this open software layer along with the associated policies and rules.

It is sought that both access to datasets, as well as services and resources on said datasets, be transparent and simple for the participants of the ecosystem, although these would not necessarily be consolidated, but rather each one resides in their own computer systems and with its own characteristics. Precisely, the work to be carried out by GAIA-X consists first of defining a process architecture that organizes such a federation of computer systems, addressing how to guarantee the correct identification of both providers (of data and services) and users, as well as the minimum set of rules (technical and governance) that must be met to enable such a federated model.

GAIA-X (Figure 5) is a particularly relevant initiative regarding Digital Sovereignty, since it aspires to create coordination between the data layer (where data sets are shared, transformed, and exploited; in the image below, it is the blue layer) and the of infrastructure and resources (in red). This coordination between the layers is the key, since GAIA-X does not seek to create projects based on data (there are other European initiatives with this objective), nor does it operate cloud services (it is not the "European cloud" in the sense of competing against offers already available on the market).

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 101037084.

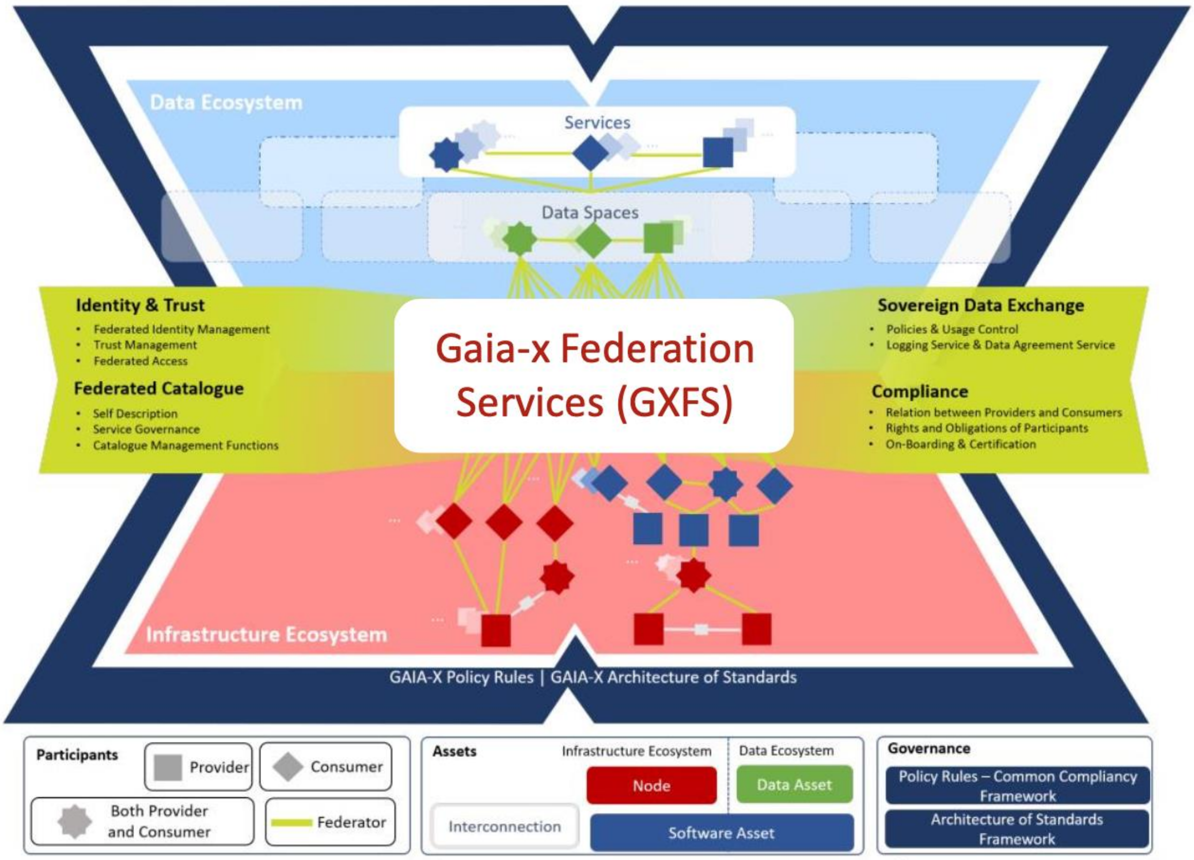

Figure 5. GAIA-X Architecture Design

GAIA-X is structured through the association (currently made up of more than 350 organizations), a series of national nodes (more than 15 hubs), sectoral working groups and a development community, coordinated by a Technical Committee, and supported by technical staff (GAIA-X Lab), large tractor projects (such as the German and French GXFS projects, GAIA-X federation services) and a community of developers. Hence, GAIA-X has mainly produced a reference architecture, a description of the trust environment (Trust Framework)[1] and a document of policies and associated rules.

2.1.1.3 SIMPL

SIMPL (Figure 6) is an Initiative of the European Commission for the development of an intelligent middleware for the implementation of data clouds and federated services in Europe, which will support the big data initiatives financed by the European Commission, such as the common European data spaces. Through a tender executed by Deloitte, SIMPL has carried out a preparatory definition work, whose main results have been the following: the business case model, a market analysis, the business requirements, service level, systems information and technology, an architecture document and an implementation roadmap.

[1] Trust framework of GAIA-X foresees verifiable credentials and linked data representations as cornerstone of its future operations.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

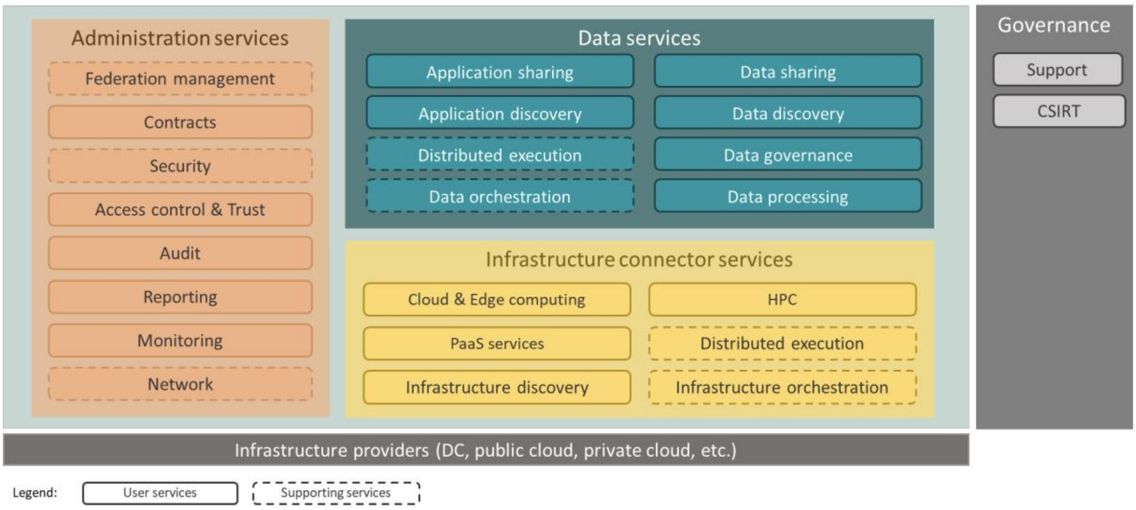

Figure 6. SIMPL architecture design and high-level architecture

The implementation plan includes the development of a minimum viable platform by the end of 2023, the launch throughout 2023 of an open test environment for experimentation, the progressive incorporation of use cases that use SIMPL, and the creation of an open-source project for the future sustainability of the platform, with the intention of producing new versions every 6 months.

2.1.1.4 Open DEI

European H2020 project for the creation of common data platforms based on a unified architecture and an established standard, with four priority sectors (manufacturing, agriculture, energy, and health) for the deployment of the EU digitalization strategy. OPEN DEI (Figure 7) has worked to detect gaps, foster synergies, support regional and national cooperation, and improve communication between innovation actions that implement the EU's digital transformation strategy. Its main result has been a position paper on data spaces ("Design Principles for Data Spaces") that has become a reference document for the main European initiatives on data sharing.

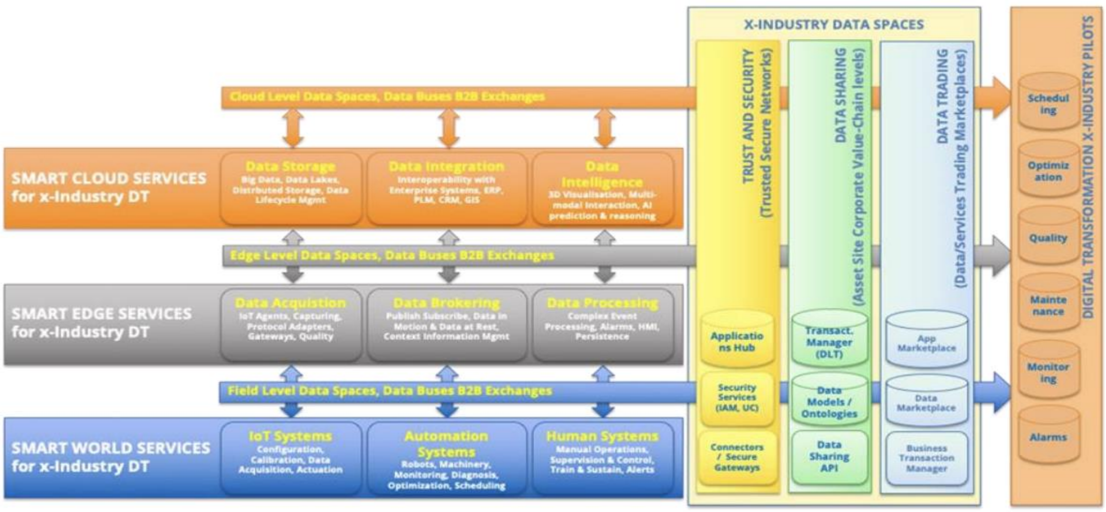

Figure 7. Open DEI general architecture

2.1.1.5 DSBA / DSSC

The Data Space Business Alliance (DSBA) is a joint initiative of BDVA, FIWARE Foundation, Gaia-X AISBL and IDSA to converge and offer a single window to all the actors involved in the data economy

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

in general. Together they represent more than 1,000 key industry players, associations, research organisations, innovators, and policy makers around the world. With this cross-sector expertise, resources, and expertise, the DSBA raises awareness, evangelizes technology, shapes standards, and enables integration across industries.

As part of this work, the DSBA is leading the consortium in charge of implementing the Data Space Support Centre (DSSC) an instrument created under the Digital Europe program as a support centre to coordinate and harmonize all relevant actions in the sectoral data spaces. The DSSC will need to enable data infrastructure architectures and requirements for data spaces, including potential technologies, processes, standards, and tools that will enable cross-sector data reuse by the public sector and European companies, in particular SMEs.

In terms of data spaces, this initiative is more focused on creating different hubs of data spaces in an alliance and a framework to share data at large scale (Figure 8).

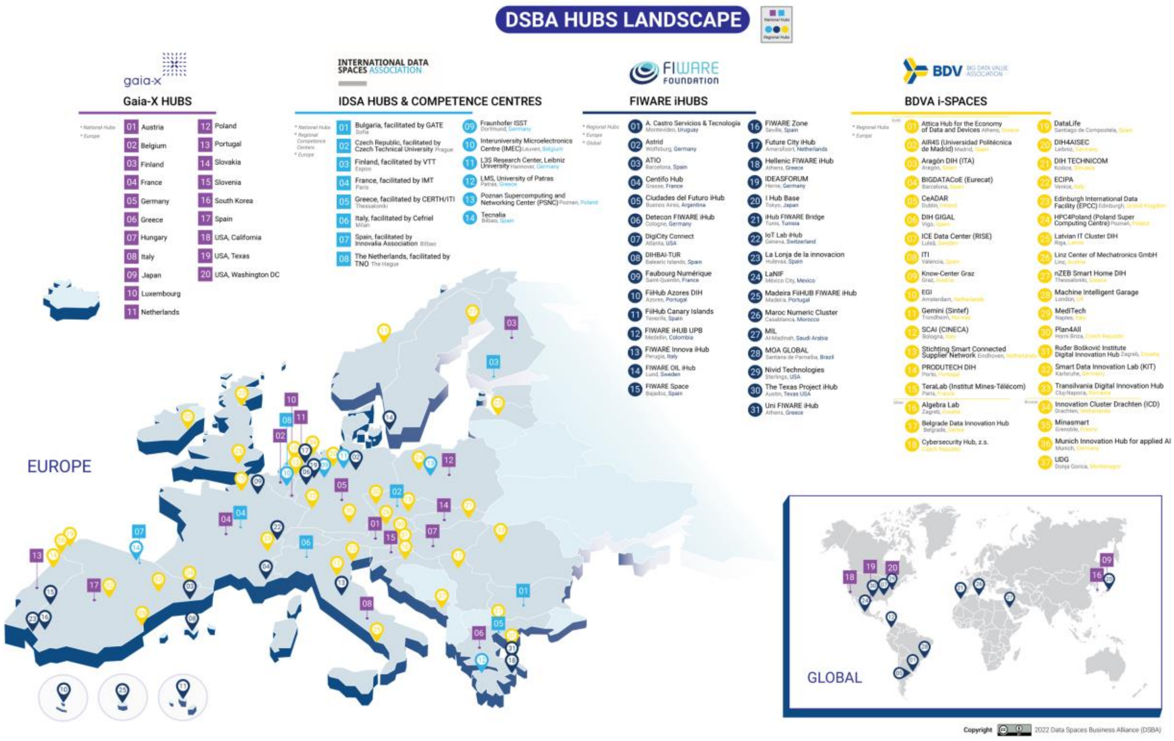

Figure 8. DSBA/DSCC open data space and centres

2.1.1.6 FIWARE

FIWARE brings a curated framework of key Open-Source software (Figure 9) components and standards which enable the development of interoperable smart solutions based on data, as well as their exchange and reuse in advanced technologies, such as IoT. FIWARE reference architecture enables plug-and-play modules for the adaptation of the platform to different domains and industries. In this regards, FIWARE architecture is mainly sustained in the ETSI-CIM (NGSI-LD) standard that enables a common representation of the information compatible with JSON-LD and Linked-Data principles.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

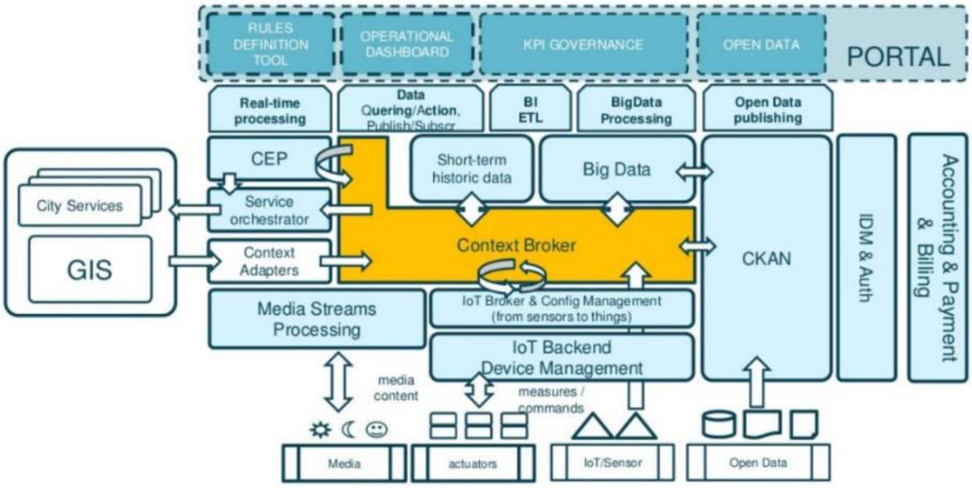

Figure 9. FIWARE Reference Architecture

This aspect permits to elaboration solutions that can integrate information from different systems, repositories, and architectures (called southbound API based on Agents). Also, the architecture enables to link the data with other third-party components/application to create full-fledged systems (Northbound API).

A part of this, FIWARE created a community around the elaboration of reusable and smart data models to represent entities and processes from different domains. This aspect permits to maintain an active community of people that permits to test and validate the data models in real scenarios.

Therefore, the main features of the FIWARE architecture relies on: (i) Modular application that could be transferable and replicable at different scales; (ii) Management of context information in a highly decentralised and large-scale manner; (iii) Open-source implementation and compatibility with informational standards such as NGSI-LD.

2.1.1.7 iSHARE

iSHARE is the European standard trust network for international business data sharing (Figure 10). The iSHARE open data space is focused on providing (i) a federated legal framework for all participants; (ii) trust registration and administration; (iii) discovery data and catalogues across the network; (iv) trust and governance. To enable this secure data space, all participants sign the same NDA and Terms of Use, and data sharing adhers to iSHARE's technical specifications. Each party is validated through a rigorous registration process.

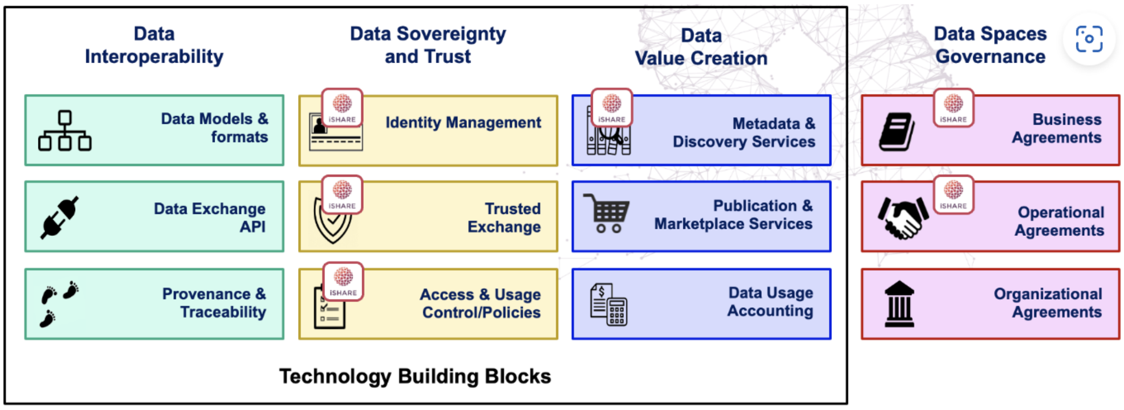

Figure 10. iSHARE open data space

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

2.1.2 Common view and comparison between Open Data Spaces Initiatives

Considering the descriptions provided in the previous section, this part of the document is devoted to providing a general overview of the Open Data Spaces highlighting their commonalities and difference (Table 1).

Table 1. Comparison between Open Data Spaces

Open Data Space	Architecture	Standards used	Security and Trust	Data Interoperability
IDSA	Peer-to-peer		Roles; Identity Management; Certification	Vocabularies/Catalogues/Connectors
GAIA-X	Federated Platform	ISO/IEC 24760-1	ISO 27001; Trust Management Module	Catalogue Management Function
SIMPL	Federated & Modular Platform		Access Control & Trust; Encryption; Authenticity and Integrity	Catalogues & Vocabulary modules
Open DEI	Cloud-based platform		Trust & Security Module	Data Models & Ontologies
DSBA/ DSCC	Hub			
FIWARE	Distributed & Modular platform	ETSI-CIM	IDM& Auth	Smart Data Models and Context brokers
iSHARE	Distributed Ledger architecture		Identity Management; Control policies	Data Models

Considering this information, the main commonalities and interoperability between open data spaces are the following ones:

1. **IDSA** Data space seems to be the standard the facto reference architecture.
2. There exists **mapping of IDSA (IDS) components with GAIA-X** components.
3. **FIWARE** is the **true/real connector** with high compatibility **with IDSA architecture**.
4. **FIWARE context broker is compatible with iShare architecture.**

2.1.3 Future directions and challenges in Open Data Spaces

Based on the overview of the data spaces and despite there already exist several implementations and designs, the Table 2 shows the main challenges and future directions of the open data spaces.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

Table 2. Future direction of Open Data Spaces

Pillars	Challenge	Description
Organizational	**Motivation & trust**	Lack of trust and motivation of organizations to share their proprietary data with third-parties
	Data lifecycle Management	Data sharing is not seen as an integral part of the data lifecycle management processes
	Human & Organizational Factors	Resistance to change, slow bureaucratic processes, lack of funding or investment
Technical	**Access & Usage Control**	Apply the access and usage control policies defined by the provider throughout the sharing process
	Security & Data Transmission	Guarantee the security of data transmission by using secure protocols and relevant techniques (e.g. encryption)
	Quality & Veracity	Control the quality and veracity of the data from origin to destination
	Volume & Velocity	Develop techniques able to process vast amounts of data in (near) real-time
	Standardization & Interoperability	In both the shared (meta)data models as well as the used technical components in compliance with FAIR principles
Legal	**Personal Data Protection**	Guarantee adherence to the GDPR and other relevant regulations and laws
	Regulatory Compliance	Ensure compliance with national, local as well as enterprise regulations in terms of data sharing

To conclude, FIWARE and IDSA open data spaces seems to be more wide adopted ones in the longer term. Furthermore, FIWARE, IDSA and GAIA-X seem to have a collaboration in the design and implementation of data spaces in EU for different domains. For that reason, **we selected FIWARE compatibility to integrate and share data across different domains inside IMPETUS**.

2.2 Cross-Domain Context Brokers compatible with ETSI-CIM standard

This second part of the section presents the different FIWARE context brokers that are currently available (Section 2.2.1). With this overview, we will analyse principal future trends of such technology for the near future (Section 2.3)

2.2.1 FIWARE Compatible Context brokers

The following context brokers \are public, and FIWARE -compliant, and published under the “FIWARE Generic Enablers”. Therefore, they can be integrated as part of any platform that adheres to the “Powered by FIWARE” statement. FIWARE is a curated framework of open-source platform components which can be assembled with other third-party platform components to accelerate the development of Smart Solutions.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 101037084.

The context broker (see Figure 11) is a piece of software aimed at the acquisition and publication of information. This software tool contains two main parts: The context Producers and the Contest Consumers. The **Context Producers** corresponds to external entities that generates information to be stored or integrated (e.g., sensors, data repositories, etc). On the other hand, there exist the **Context Consumers** that corresponds to entities interested in consuming the information (e.g., applications, analytical services, other architectures, etc). Complementary, the exchanged information is considered as events that can be handled by applications or platform components which subscribe to those events.

The context broker supports two-way communications: push or pull information towards both the Context Producers and the Context Consumers. The Context Producer may continuously push the context information into the Context Broker. The Context Broker, respectively, can request the context information from Context Producers if they provide the ability to be queried. In a similar way, Context Consumers can pull the context information from the Context Broker, while the Context Broker can notify on updates in context-to-Context Consumers interested in such updates by means of subscriptions.

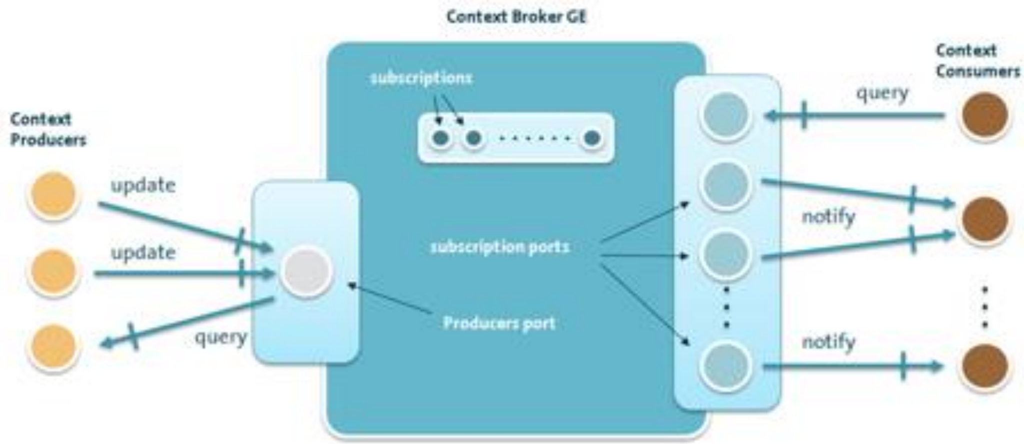

Figure 11. General architecture of the context broker

2.2.2 Orion-LD Context Broker

[Orion-LD is a Context Broker] and CEF building block for context data management which supports both the NGSI-LD and the NGSI-v2 APIs. It is currently a fork of the original Orion Context Broker extending support to add NGSI-LD and linked data concepts. Orion-LD follows the ETSI specification for NGSI-LD and has been tested to be a stable and fast NGSI-LD broker with close compliance to the version 1.3.1 of the NGSI-LD API specification.

Orion Context Broker stores context information updated from applications, so queries are resolved based on that information. Users can do the following operations by using the Orion interface: query context information, update context information, provide notifications about changes on context information, and finally, register context provider applications.

In terms of architecture (Figure 12), the Orion Context broker is formed by the southbound API in charge of managing IoT Agents subscriptions and store the information into a MongoDB database. On the other hand, there exist the northbound API in charge of gathering the requests and subscriptions from the consumers and provide to them batch or stream information. The Orion-LD engine is in charge and balance such subscriptions to store and retrieve information accordingly.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

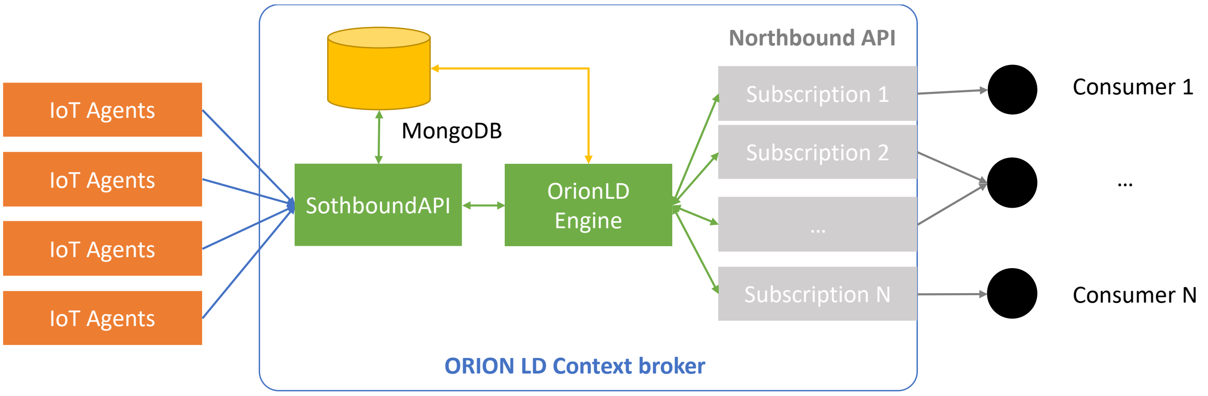

Figure 12. Orion Context Broker architecture

2.2.3 Scorpio Broker

Scorpio is an NGSI-LD compliant context broker developed by NEC Laboratories Europe and NEC Technologies India. It implements the full NGSI-LD API as specified by the ETSI Industry Specification Group on cross cutting Context Information Management (ETSI ISG CIM). Scorpio's functionalities include the management of the context information (create, update, modify and delete), query it, including filtering, geographic scoping and paging, subscriptions to context changes and the receiving of asynchronous notifications. As an addition, Scorpio also includes the register and discovery of sources of context information, allowing building distributed and federated deployments.

In terms of architecture (Figure 13), Scorpio context broker manages the internal communication and IoT agents streaming information using Kafka. Scorpio collects these events and store the collected and fused information inside a PostgreSQL database. At the Northbound, Scorpio provides a request/response Open API constructed using Swagger under Spring environment.

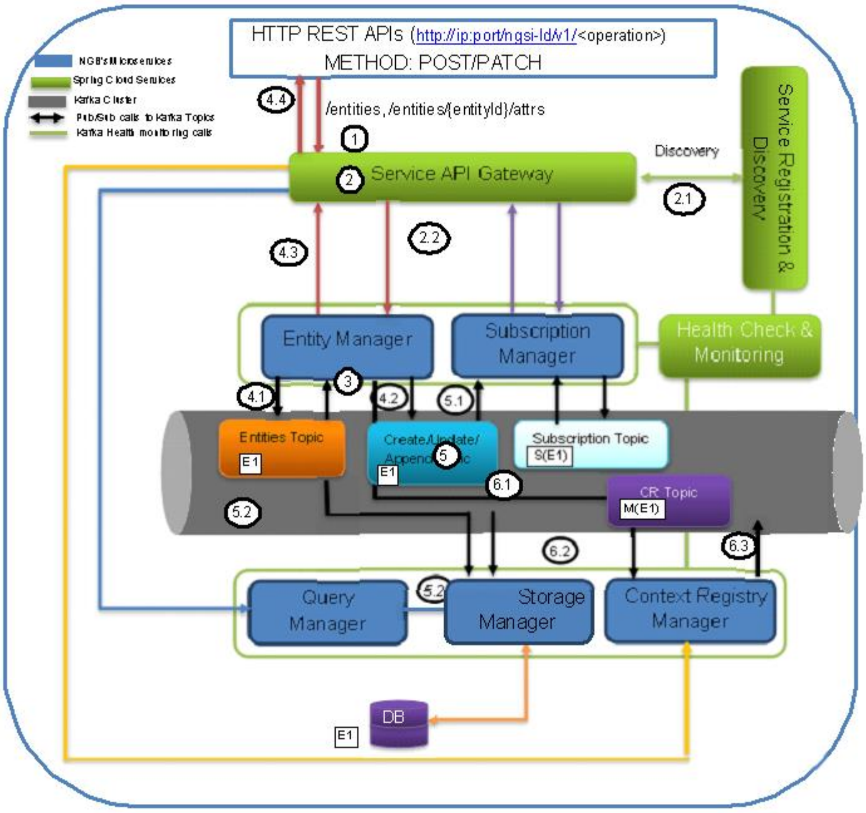

Figure 13. Scorpio Context Broker Architecture

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

2.2.4 Stellio Context Broker

Stellio is an NGSI-LD compliant context broker developed by EGM based on Spring Boot Framework (Figure 14), developed in Kotlin, and built with Gradle. It provides an API Gateway module that dispatches requests to downstream services as well as a Kafka streaming engine that decouples communication inside the broker and allows plugging other services seamlessly. It provides an Entity, Search and Subscription service in charge of managing the information context, handling the temporal and geospatial queries, and managing subscriptions and subsequent notifications respectively.

In terms of data storage, Stellio stores the information using Neo4j for the entity's description. Also, it uses PostgreSQL for storing time-series related information.

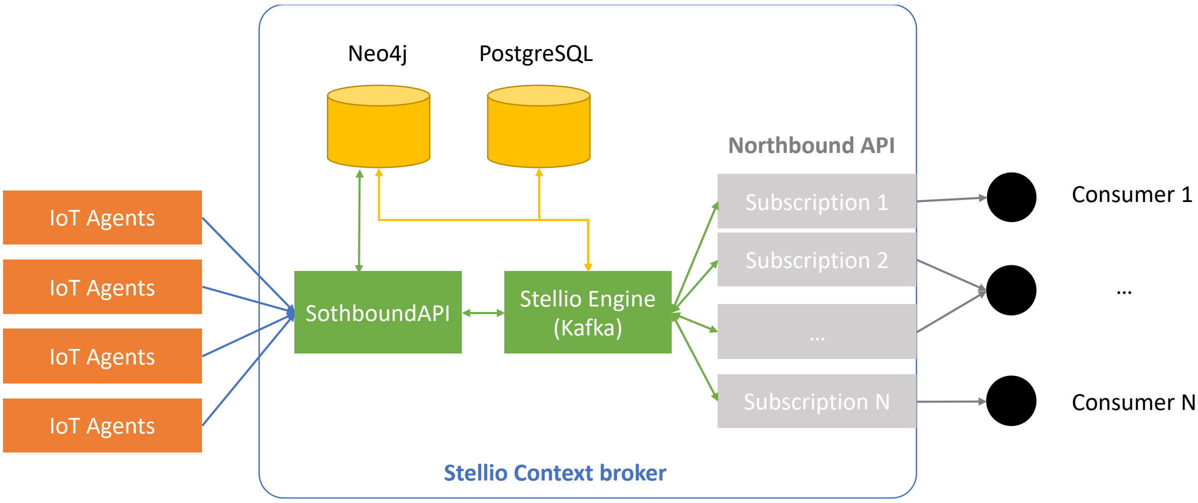

Figure 14. Stellio general architecture

2.3 Analysis of context brokers and future directions

Considering the general information about the context brokers, the main summary and overall information is described in the Table 3.

Table 3. Comparison between FIWARE based context brokers

Context Broker	Technology	Databases	Event Engine
Orion	C based platform	MongoDB	
Scorpio	Java/Spring	PostgreSQL	Kafka
Stellio	Java/Spring	Neo4J/PostgreSQL	Kafka

Based on this information, the main challenges for the context brokers are summarised in the Table 4.

Table 4. Context broker's main challenges

Pillars	Challenge	Description
Organizational	**Motivation & trust**	Lack of trust and motivation of organizations to share their proprietary data with third parties

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

	Human & Organizational Factors	Resistance to change, slow bureaucratic processes, lack of funding or investment
Technical	**Semantic Reasoning**	Enable semantic reasoning over the information to facilitate the induction of cross-domain events
	Split context information from Time Series	Split the storage between knowledge-graphs (for entity description and metadata) and time series medium term data.
	Manage FAIR data cycle	Enable and manage FAIR data principles for data catalogues stored in the application.
	Explainability	Enable explainability to understand events at cross-domain
	Standardization	Be compliant with standard vocabularies and standard ontologies to maximise the adoption of smart data models

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

3 Platform architecture overview

Under this chapter, we will explain the main RKB architecture designed and presented under the D2.1 ((UiT), 2022)(Section 3.1). Moreover, we will show the interrelationship between the different work packages in the construction of such architecture (Section 3.2).

3.1 RKB Architecture

The RKB architecture is envisioned as a distributed platform to collect data from multiple sources of information coming from multiple domains. During the project, we will analyse how to federate part of the platform to enable fully interaction and data sharing across regions, giving a single point to explore and consume information.

The envisioned capabilities of the platform are based on research on the main challenges described for Open Data Spaces (Section 2.1) and the Context Brokers (Section 2.2):

1. Provide the **integration of heterogeneous information harmonised** following up ontologies and **common vocabularies** to enable machine-to-machine interoperability.
2. Facilitate **data fusion using meta-data descriptions (ontology)** to understand climatic events and information from different domains.
3. Generates datasets **compatible with FAIR principles enabling** corresponding pipelines for data quality, data wrangling, and non-discrimination.
4. **Enabling security mechanism** to support trustworthy AI.
5. **Contribute to open science and open linked data** through the exposition of open data from different regional, national and transboundary information.

According to these main capabilities Figure 16 depicts the different layers in which is structured the RKB platform, each layer has one responsibility indicated as follows:

Layer	Description
Physical Layer	This layer is mainly devoted to the data acquisition from regional, national, and transboundary data repositories from different domains. The collected information will consider long-term climatic variables to sustain the different predictions and analysis. Also, this layer will collect information from in-situ digital tools such as decision-support tools, early warning systems, modelling & simulation tools and Digital Twins installed locally for specific purposes.
Data Layer	This part of the architecture is focused on managing the information according to the FAIR principles. In this regard, the context brokers will be the main element to collect and store short-term data. The proposed context broker (FIWARE compatible) will combine knowledge graphs (with semantic reasoning) with time-series databases to manipulate large volumes of information. Complementary, data quality and non-discrimination algorithms will facilitate the elaboration of data-driven algorithms.
Analysis Layer	Provides domain- and problem-specific service modules and processing tools. In this regard, this layer will provide specific data-driven algorithms to detect climatic event patterns, analysis of risks and detection of hot-spots, climatic downscaling and projections and analysis of anthropogenic factors impacting climate events.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

Visualization/Application Layer	This part of the architecture is the closest to the users providing the needed mechanism to interact and explore the information. Thus, this part of the architecture comprises the elaboration of scenarios, data visualization and the generation of custom pipelines to analyse and understand information.

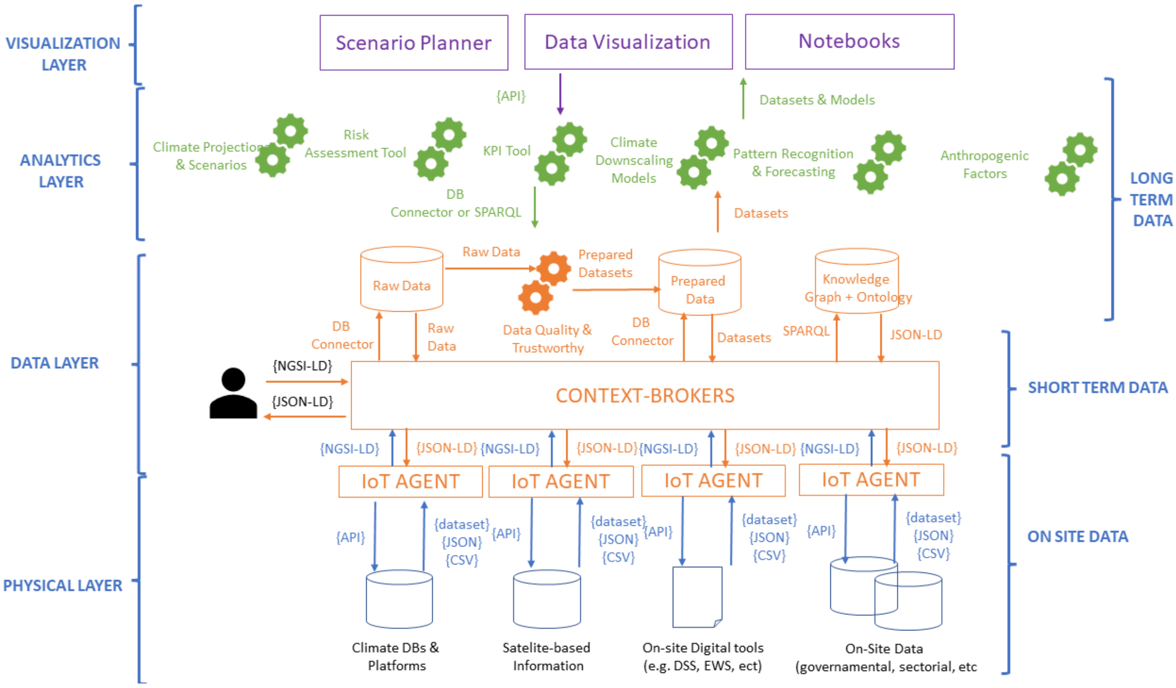

Figure 15 RKB General Architecture Schema

3.2 Interrelation of WPs in the elaboration of the virtual RKB

For the development of the envisioned platform, the collaboration of multiple technical and non-technical work packages is required. In that sense, non-technical WPs such as stakeholder engagement (**WP1**) will support the development by motivating the case-studies in the activities of data collection and visual tool development (visualization engine).

The core modules of the RKB will be elaborated under the **WP2** that will include the elaboration of the RKB structure, the implementation of the semantic enhanced reasoning (metadata-reasoning), the pipelines for ensuring data quality according to FAIR principles and data-driven algorithms for (i) pattern detection over multi-scale time series; and (ii) application of deep learning to understand anthropogenic factors.

WP3 will be focus on climate risk and adaptation modelling considering (i) climate downscaling models and projections; (ii) Risk assessment and identification of hot-spots; and (iii) the stress testing platform for generating adaptation pathways. Moreover, WP3 will provide the list of KPIs to better understand and present the information to the case-studies.

WP4 will be devoted to providing information from case-studies at two different scales. In the first scale will help in data identification at regional level coming from public data repositories in each of the regionalities. At second stage level, this WP will provide information coming from the digital assets to be developed at regional scale (e.g., early warning systems, decision support tools, digital twins, etc).

WP5 will mobilise the community to facilitate the elaboration of the policy-based strategies and corresponding interventions to ensure sustainable strategies at long-term. Indeed, this WP will support the platform interaction to understand the data and drive the guidance to elaborate the engine to determine policies understanding at different time-frames (simulation over the data).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

Finally, **WP6** and **WP7** will support the RKB understanding to the public and, the determination of the business behind the implementation of the RKBs.

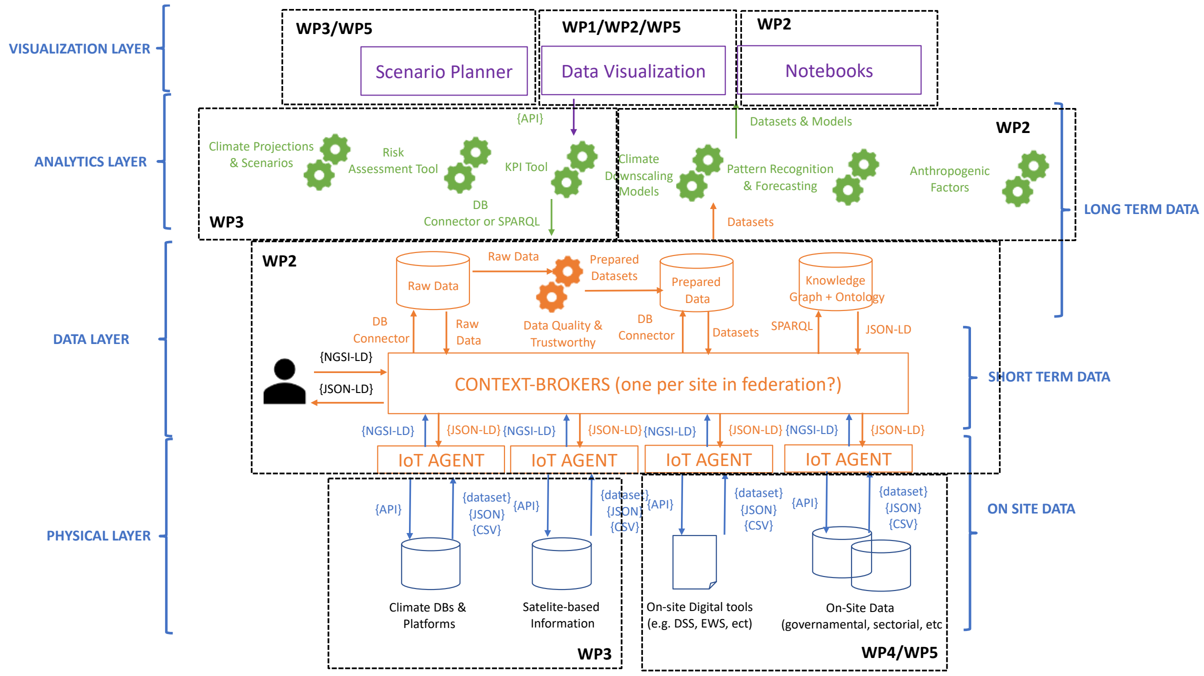

Figure 16. WPs interrelation in the RKB construction

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

4 Semantic Context Broker Tool

This chapter is devoted to the description of the Semantic Context Broker tool that will be used in IMPETUS to collect and expose heterogeneous public and private information to enable policies generation at long term. For that purpose, this section will initially describe the context broker architecture (Section 4.1). Following up the architecture description, the next section is focused on describing the construction of the semantic context broker tool (Section 4.2).

4.1 General Architecture of the context broker

This initial part of the document is mainly focused on the description of the Semantic context broker architecture (Figure 17). This tool will be FIWARE-compatible due to FIWARE technology combined with IDSA, which will probably be the de facto standard for future Open Data Spaces in Europe. The proposed context broker uses Orion-LD base with some extra capabilities in order to avoid the main barriers of existing context brokers in terms of:

- **Time consuming in data integration or consumption** when huge amount of information is integrated into the context broker.
- **Full semantic reasoning** inside knowledge graphs to understand information linkage and potential simple causal events in climate using inductive and probabilistic reasoning.
- **Stream processing** coming from repositories, or the joining integration actions from the IoT Agents.
- **Enabling semantic rules** for facilitating mappings coming from information in different types (enabling R2RML rules for data integration)
- **Enabling rules for data validation** using, for example, SHACL rules.

Considering these aspects, the Semantic Context Broker **is responsible for managing the data lifecycle according to the FAIR principles using contextual information gathered or produced from heterogeneous sources.** The main intention is to collect and produce actionable information for enabling further processing, analysis, and extraction of new knowledge.

Based on this main purpose, the envisioned architecture (Figure 17) will be composed by the following modules:

Table 5. Semantic Context Broker modules

Module	Description
IoT Agents	FIWARE compatible IoT agents to the data capturing from different data sources. In IMPETUS it will be elaborated corresponding needed IoT agents if necessary to capture data from public and private repositories.
Southbound API	This module will capture the data streams from the IoT Agents and to ensure non-blocking information in the platform, a balancer will be applied to capture the information as an event to be managed by the Stream manager.
Stream Manager	Module responsible to capture the different tasks (events) occurring inside the context broker (entity publication, data gathering, answering request, resolving subscriptions, etc). Then this module will manage all these events to produce different responses and manage the complete flow of information.
Knowledge graph	A graph database to store contextual information about the entities.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

Time Series DB	A database to store the different time series of data coming from the different public and private repositories
Geospatial DB	A geospatial database to store the different geospatial layers and information.
Northbound API	The Open API that will allow the connection of different consumers into the platform. These connections could be of request/response type (to receive static information) or publication/subscription type (stream or real-time information with regards to the time it is ingested by the context broker).

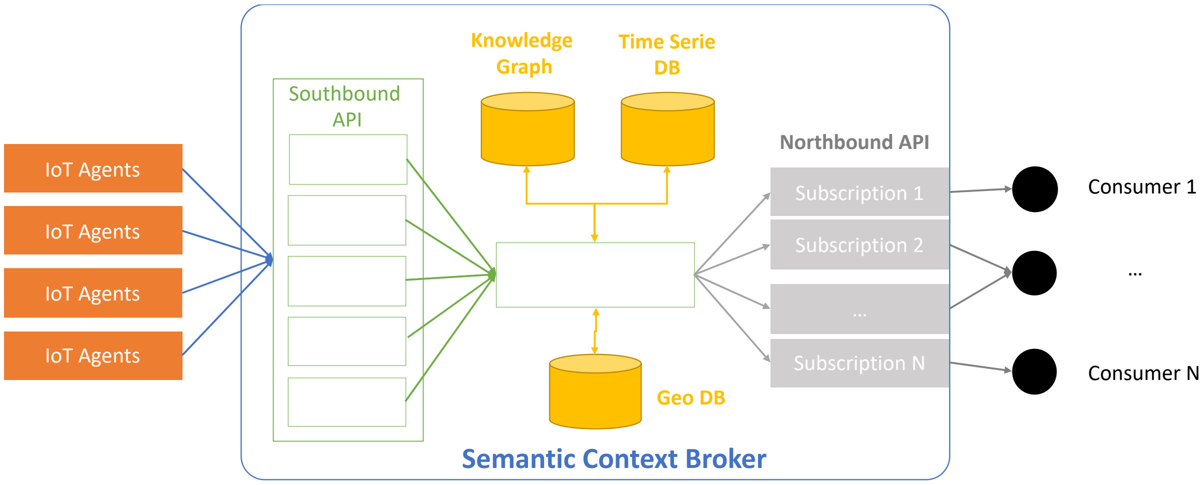

Figure 17. Semantic Context broker Architecture

4.2 Construction of the Semantic Context Broker tool

This specific section describes the components of the semantic context broker. Hence, this section starts with the definition and specification of the used technology (Section 4.2.1). After that, we will explain the main properties of the data models (called "Context") (Section 4.2.2). One we have a clearer vision about the data models, we will explain the southbound (Section 4.2.3) and the Northbound (Section 4.2.4) APIs. We will finalise this section by exposing the public and private data that currently we expect to integrate (Section 4.2.5).

4.2.1 Technology used in the Semantic Context Broker Tool

The Table 6 describes the main technology used for the semantic context broker in relation to the architecture described in the previous section. The development environment for the entire platform structure will be Python-based.

Table 6. Technology used in the semantic context broker

Module	Technology
IoT Agents	FIWARE IoT Agents
Southbound API	Kafka
Stream Manager	Kafka Streams or Spark Streaming

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

Knowledge graph	Stardog/Jena/Proxygraph
Time Series DB	Druid/TimeSeriesDB
Geospatial DB	PostgreSQL/MongoDB
Northbound API	Swagger

4.2.2 Semantic Context Broker: The Context

The **context** refers to the generic data model structure that is the base to describe the entities as a whole (e.g. repositories, decision-making tools, etc). The data models structure corresponds to a JSON-LD[2] format, a serialization of the NGSI-LD[3] specification, which not only defines the foundational concepts (Entities, Relationships, Properties), but also an API for publishing, querying, and subscribing to context information, facilitating the sharing of information.

During this specific part of the document, we will describe: (i) the NGSI-LD Entity (Section 4.2.2.1), (ii) the NGSI-LD Property (Section 4.2.2.2) and (iii) the NGSI-LD Relationship (Section 4.2.2.3).

4.2.2.1 NGSI-LD Entity

The NGSI-LD Information Model prescribes the structure of context information that shall be supported by an NGSI-LD system. It specifies the data representation mechanisms that shall be used by the NGSI-LD API itself. In addition, it specifies the structure of the Context Information Management vocabularies to be used in conjunction with the API.

The JSON-LD object shall contain at least the following members to mainly describe an Entity. It is worthy to remember that an entity corresponds to all representative data generators (e.g., a data repository, a decision-making tool, early warning system, etc). The main attributes defined for an entity are:

Table 7. Definition of a NGSI-LD Entity

Attribute	Type	Description
id	URI (Uniform Resource Identifier)	URI with which an entity can be recognized
type	NGSI-LD Entity	Type of the entity (e.g. WeatherEntity, etc)
@context	URI or List<URI>	URI or list of URIs that corresponds to the potential vocabularies to be considered in the schema
property	NGSI-LD Property	Array of properties (e.g., temperature) that are measured by the entity. In case of multiple Property instances with the same Property Name, all instances are provided as an unordered JSON array

[2] https://json-ld.org/primer/latest/
[3] https://www.etsi.org/deliver/etsi_gs/CIM/001_099/009/01.01.01_60/gs_CIM009v010101p.pdf

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

relationship	NGSI-LD Relationship	Array of Relationships that represent the link between the different entities (e.g., a Weather Repository that is related with an Entity Regional Temperature)

4.2.2.2 NGSI-LD Property

An NGSI-LD Property corresponds to the measurements performed by a defined entity. In this regard, a NGSI-LD property shall be represented by a member whose key is the Property Name and whose value is a JSON-LD object (or JSON-LD array with such JSON-LD objects if there are multiple instances with the same Property Name). The attributes that define a property are the ones represented in the Table 8.

Table 8. NGSI-LD Property attributes

Attribute	Type	Description
type	URI	Type of the property. It is a mandatory field
value	Object	It is a mandatory field that represents the specific value of the measurement. It contains the attribute “@type” as an URI to the specific data type and the attribute “@value” to represent the corresponding measurement.
object	URI	Normally it is not used, and it represents the NGSI-LD relationship object.
observerdAt	String	Mandatory field that represents the data in which the property is observed.
datasetId	URI	Mandatory field that represents the URI in which this measurement is present
instanceId	URI	Mandatory field that represents the URI of the NGSI-LD property
createdAt	String	String that represents the date in which the property is created.
modifiedAt	String	String that represents the date in which the property is modified (last modification).
unitCode	String	It corresponds to an optional parameter. The “*unitCode*” is represented as a String that represents the measurement unit corresponding to the Property value. It shall be encoded using the UN/CEFACT Common Codes for Units of Measurement (United Nations, 2005).
properties	NGSI-LD Property	Optional field, a property itself can be an attribute, it can be nested in a property by defining a key, the property name and a value inside.
relationships	NGSI-LD Relationship	Optional field value that contains information on how an entity (or an attribute) relate to other entities. A relationship must have an "object", and the object must be a string that is a URI.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 101037084.

4.2.2.3 NGSI-LD Relationship

NGSI-LD Relationship corresponds with the interlink between entities or between entities and properties. The NGSI-LD relationships are represented as a JSON-LD objects that contains the following attributes:

Table 9. NGSI-LD Relationship attributes

Attribute	Type	Description
type	URI	Type of the property. It is a mandatory field
value	Object	It is a mandatory field that represents the specific value of the measurement. It contains the attribute “@type” as an URI to the specific data type and the attribute “@value” to represent the corresponding measurement.
object	URI	Normally it is not used, and it represents the NGSI-LD relationship object.
observerdAt	String	Mandatory field that represents the data in which the property is observed.
datasetId	URI	Mandatory field that represents the URI in which this measurement is present
instanceId	URI	Mandatory field that represents the URI of the NGSI-LD property
createdAt	String	String that represents the date in which the property is created.
modifiedAt	String	String that represents the date in which the property is modified (last modification).
relationships	NGSI-LD Relationship	Optional field value that contains information on how an entity (or an attribute) relate to other entities. A relationship must have an "object", and the object must be a string that is a URI.

4.2.2.4 Example of a NGSI-LD Schema

To depict the Semantic context broker data models, the Listing 1 represents an extraction of an observation of air quality conditions at a certain place and time. The whole JSON schema is described in the repository of Fiware data-model. The fields included in the schema are explained in Table 10.

Table 10. Air quality schema explanation

Attribute	Type	Description
type	URI	Entity type, AirQualityObserved
dateObserved	Object	The date and time of this observation in ISO8601 UTCformat.
NO2	URI	Property defining the quantity of Nitrogen dioxide found in the observation

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 101037084.

refPointOfInterest	String	A reference to the device(s) which captured this observation.
@context	URI	URIs where all the fields for an AirQualityObserved entity are defined.

```
{
    "id": "urn:ngsi-ld:AirQualityObserved:RZ:Obsv4567",
    "type": "AirQualityObserved",
    "dateObserved": {
        "type": "Property",
        "value": {
            "@type": "DateTime",
            "@value": "2018-08-07T12:00:00Z"
        }
    },
    "NO2": {
        "type": "Property",
        "value": 22,
        "unitCode": "GP"
    },
    "refPointOfInterest": {
        "type": "Relationship",
        "object": "urn:ngsi-ld:PointOfInterest:RZ:MainSquare"
    },
    "@context": [
        "https://schema.lab.fiware.org/ld/context",
        "https://uri.etsi.org/ngsi-ld/v1/ngsi-ld-core-context.jsonld"
    ]
}
```

Listing 1. NGSI-LD Example for an observation of air quality conditions

4.2.3 The Southbound API: The Pub/Sub Architecture (Publication)

As described at the start of the section 4.2.2, NGSI-LD also specifies a publication/subscription mechanism for injecting data into the context broker. The publication mechanism is done by means of the operations defined in the context broker REST-API (Table 12).

Thus, the different publishers can use the following post operation POST operation as described in the following table (Table 11):

Table 11. Southbound API operations

URL	Operation	Input	Output	Description
/entities	POST	NGSI-LD Entity	NGSI-LD Entity	Creation of single entity into the context broker
/entityOperations/create	POST	List< NGSI-LD Entity>	List< NGSI-LD Entity>	Creation of multiple entities into the context broker. An error is thrown if the entity already exists.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

/entityOperations/upsert	POST	List< NGSI-LD Entity>	List< NGSI-LD Entity>	Creation of multiple entities into the context broker. Updates the entity with the provided values if already exists.
/entities/{entityId}/attrs	POST	NGSI-LD properties	NGSI-LD Entity	Creation of properties inside an entity.
/entities/{entityId}/attrs/{attrName}	PATCH	NGSI-LD properties	NGSI-LD properties	Update a specific property inside an entity.

As depicted in Figure 18, a provider can publish data by means of creation or modification of the information that is collecting. Using the described operations, the provider could store the data in the context broker database for further usage.

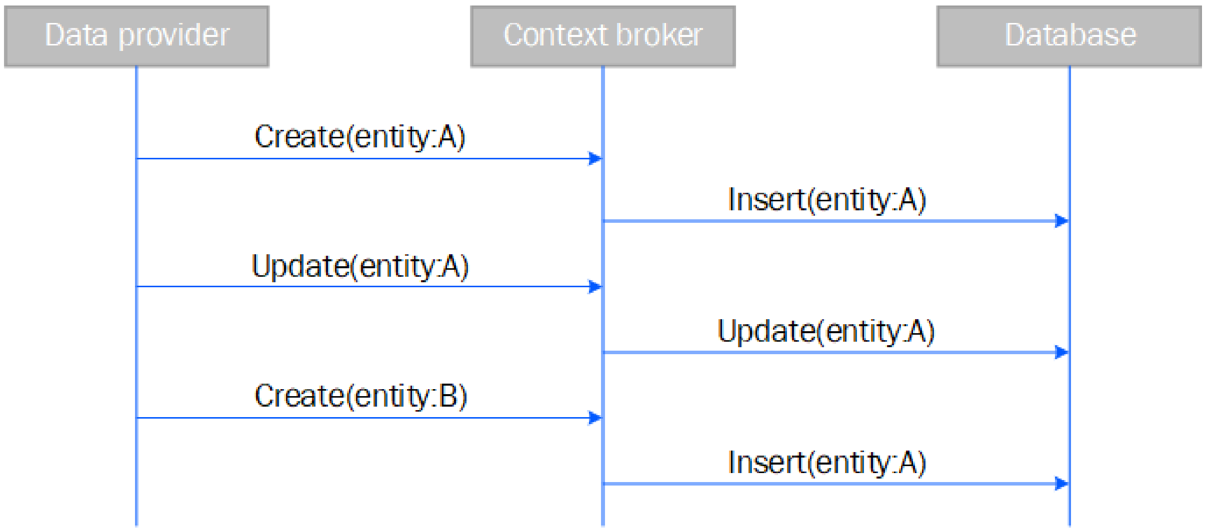

Figure 18 Context broker publish mechanism

4.2.4 The Northbound API

As mentioned in the Semantic context broker architecture, the context broker offers the Northbound API to expose data to the consumers. This data exposition could be performed using two main methods: (i) The Open API (Section 4.2.4.1) and the Pub/Sub (Subscription architecture) (Section 4.2.4.2)

4.2.4.1 Open API

The NGSI-LD Open REST API corresponds to the request/response access to the information contained inside the context broker. Indeed, this API exposes the main endpoints for manage the context information from the most basic operations creating, retrieving, updating, and deleting entities, as well as a subscription mechanism for receiving updates of the information in real time.

4.2.4.1.1 Entity

The operations defined under this section refer to all needed actions to manage an entity inside the context broker (Table 12).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

Table 12. Definition of the "Entities" operation for the REST-API

URL	Operation	Input	Output	Description
/entities	POST	NGSI-LD Entity	NGSI-LD Entity	Creation of single entity into the context broker
/entityOperations/create	POST	List< NGSI-LD Entity>	List< NGSI-LD Entity>	Creation of multiple entities into the context broker. An error is thrown if the entity already exists.
/entityOperations/upsert	POST	List< NGSI-LD Entity>	List< NGSI-LD Entity>	Creation of multiple entities into the context broker. Updates the entity with the provided values if already exists.
/entities/{entityId}	GET		NGSI-LD Entity	Obtain a specific entity identified by an id.
/entities?id=<id1, id2, …>	GET		List<NGSI-LD Entity>	Obtain multiple entities identified by different ids.
/entities?type{type}	GET		List<NGSI-LD Entity>	Obtain a list of entities considering their type.
/entities/{entityId}/attrs	POST	NGSI-LD properties	NGSI-LD Entity	Creation of properties inside an entity.
/entities/{entityId}/attrs	PATCH	NGSI-LD properties	NGSI-LD Entity	Modification of one or more attributes inside a NGSI-LD entity.
/entities/{entityId}/attrs/{attrName}	PATCH	NGSI-LD properties	NGSI-LD properties	Update a specific property inside an entity.
/entityOperations/delete	POST	List<NGSI-LD Entities>		Deletion of multiple entities
/entities/{entityId}	DELETE			Delete a specific Entity by Id
/entities/{entityId}/attrs/{attrName}	DELETE			Delete a specific attribute from an specific entity

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

4.2.4.1.2 Subscription Operations

The operations defined for the “subscription” operation refer to the registration and de-registrations of consumers into the context broker. The specific operations are the ones described under the Table 13.

Table 13. Operations for the "Subscription" operation of the REST-API

URL	Operation	Input	Output	Description
/subscriptions	POST	NGSI-LD Entity	NGSI-LD Entity	Creation of an NGSI-LD Entity with type Subscription into the context broker.
/subscriptions/{subscriptionId}	PATCH	NGSI-LD Entity	NGSI-LD Entity	Update of a specific subscription into the context broker.
/subscriptions	GET		List< NGSI-LD Entity>	Obtain the list of subscriptions available inside the context broker
/subscriptions/{subscriptionId}	GET		NGSI-LD Entity	Obtain the information of an specific subscription.
/subscriptions/{subscriptionId}	DELETE			Deletion of a specific subscription from the context broker.

4.2.4.2 The Subscription information gathering (Pub/Sub)

This part of the document is focused on describing the consumption part of the northbound API. That means, information gathering from the context broker using the subscription approach.

Thus, the different consumers that are subscribed to a specific entity receive updates of the information in real time, while they are changing inside the context broker (receiving information from the southbound API). As depicted in Figure 19, using this mechanism, a client can request the context broker to notify him on certain updates in the data. This is achieved using the “subscribe” operation. This operation allows the client to specify the notification channel. Moreover, the client can focus on specific data of interest by providing filters over the entity id, entity type, attribute, etc. Once subscribed, whenever a data provider updates an entity that matches the filters provided by the client in the subscription operation, the context broker will automatically notify the client of this event.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 101037084.

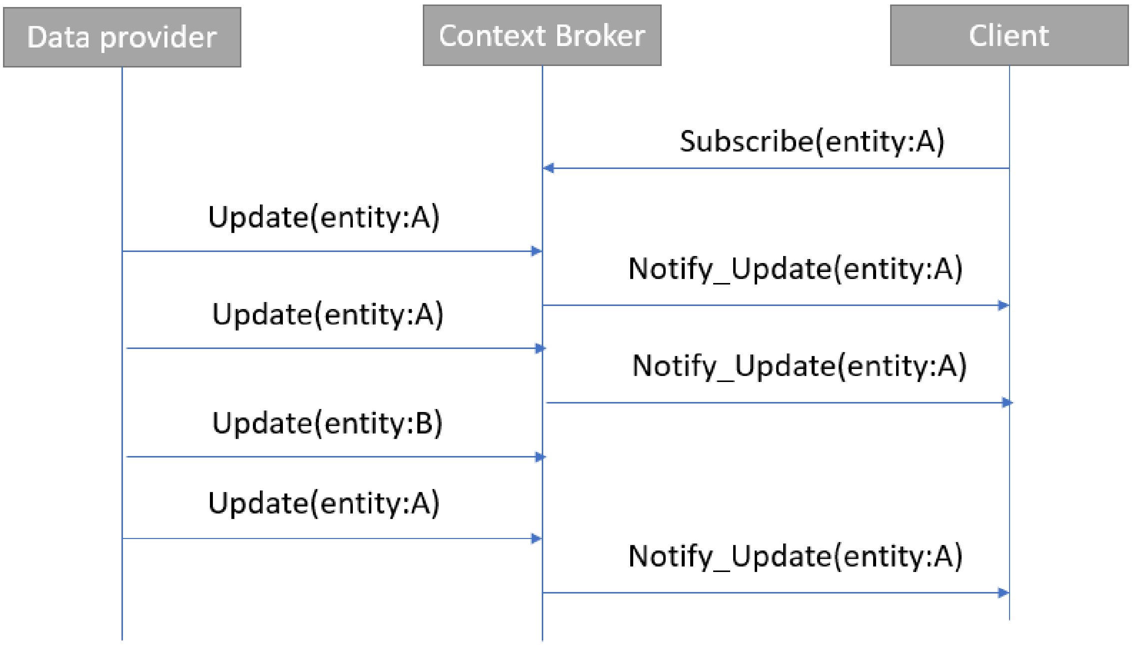

Figure 19 Context broker subscription mechanism

Considering this sequence of actions, the commonly information exchanged between the context broker and the different consumers corresponds to a JSON-LD file with the following attributes (Table 14).

Table 14. Subscription data model attributes

Attribute	Type	Description
type	URI	Type of the property. It is a mandatory field
entities	NGSI-LD Entity or List<NGSI-LD Entities>	It is a mandatory field that represents the entity or set of entities that are being changed.
WatchedAttributes	NGSI-LD Properties	Optional attribute. A list of attributes (URI) that are being watched for changes. If there is an empty list, it means the system will look at any changes inside the defined entities.
q	String	Optional attribute. Condition that corresponds to the trigger to launch the different notifications.
geoQ	String	Optional attribute. Condition that corresponds to geospatial triggers to launch the different notifications.
notification	Object	Mandatory attribute. Object with the parameters to configure the notification output
attributes	List<NGSI-LD Property>	Mandatory attribute. List indicating the attributes that must be present in the notification
format	String	Optional attribute. Output notification format
endpoint	URI	Mandatory attribute. Object indicating the URI of the notification receiver, and the data type to accept

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

uri	URI	Mandatory attribute. HTTP or MQTT URI that identifies the stream server.
accept	String	Type of the HTTP application accepted
value	String	Occasionally present and defines the property vale.

4.2.5 Public and Private Data to be integrated into the Semantic Context Broker

During this section, we are going to describe the potential data to be integrated from the public regional, national, and European repositories directly or indirectly (across other tools implemented in IMPETUS) (Section 4.2.5.1). Complementing this information, we will describe initial introspection of data gathering from the case-studies (Section 4.2.5.2).

4.2.5.1 Public Data

Public data refer to regional, national, and European data repositories that are interesting for IMPETUS project and the subsequent biogeographical regions. It is interesting to mention that the data capturing from the presented repositories in Table 15 will be gathered with the following ways:

1. **Directly**. Data captured from the URL of the repository directly through API, FTP or URI.
2. **Indirectly**. Using third-party system that already have that information (e.g., through other tools of the IMPETUS tools).

Based on this notation, the different repositories to be integrated are:

Table 15. Public Repositories identified in IMPETUS

#	Identifier/Name	Type of Access	URL
1	Processes Risk assessment from (ERA5, CMIP6, EUROSTAT, IIASA)	API	https://impetus.uwmh.eu
2	Copernicus	API	https://www.copernicus.eu/
3	ESA Data Cubes	API	https://eurodatacube.com/documentation/about
4	European Environment Agency (EEA)	API	https://www.eea.europa.eu/themes
5	Group of Earth Observation (GEOSS)	API	https://www.earthobservations.org/geoss.php
6	Water Information System for Europe (WISE)	API	https://water.europa.eu/
7	Climate Change Initiative Open Data Portal	API/HTTP	http://cci.esa.int/data
8	European Climate Adaptation Platform Climate-ADAPT	Web	https://climate-adapt.eea.europa.eu/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

9	GeoPortal del Ministerio de Medio Ambiente	API	https://www.miteco.gob.es/es/cartografia-y-sig/ide/geoportal/

4.2.5.2 Private Data

The **private data**, under this specific context, we refer to that kind of information that is not already available using the web. This data refers to specific regional tools or datasets that are or will be in internal servers from the case-studies. In this relation, the list of private data envisioned to be included through the context brokers are depicted in Table 16.

Table 16. Private data to be included through context broker

#	Identifier/Name	Responsible Partner	URL (if Exists)
7	TOOL-WP3-RISK-ASSESSMENT-MODELS	NTUA	https://impetus.uwmh.eu
9	TOOL-WP4-BIODIVERSITY & REFORESTATION TOOL	MAICH	TBD
10	TOOL-WP4-DIGITAL TWIN SEWER MINING	NTUA	TBD
11	TOOL-WP4-WATER-ENERGY-SIMULATION	NTUA	TBD
12	TOOL-WP4-DSS-IWM	KWB	TBD
13	TOOL-WP4-ONLINE-PATHOGEN MONITORING	EUT	TBD
14	TOOL-WP4-FLOOD RISK MODEL	N&S	TBD
15	TOOL-WP4-DIGITAL-TWIN-BUSINESS	NTUA	TBD
16	TOOL-WP4-EWS-FLOODS	BEF	TBD
17	TOOL-WP4-SATELLITE-COASTAL-MONITORING-SYSTEM	LOB	TBD
18	TOOL-WP4-EWS-AVALANCHES	UiT	TBD
19	TOOL-WP4-MARINE SPATIAL FRAMEWORK	TFFK	TBD
20	TOOL-WP4-DST-INDUSTRI-DECARBONIZATION	WEI	TBD

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

5 Prototype demonstration

In this chapter we present how to use the context brokers presented in Section 2.2.1, although they all are different, they all follow the NGSI-LD implementation, therefore, its usage must be similar. Furthermore, they also have the possibility to be deployed by means of Docker containers.

To demonstrate some of the functionalities of the context broker two clients have been developed in Python that interact with the specified API. Below, a description of the configuration to deploy the context broker and the operation of these sample clients is provided and screenshots of the HTTP communication between the clients and the prototype are given.

As mentioned previously, the deployment of the context broker will be done by means of a Docker container, to do so, a docker-compose.yml file is needed. This file will contain the context broker and the parameters to be able to access it as well as a Mongo database to be able to store the data produced by the context producer and recovered by the context consumer (see *Figure 11*)

```
version: "3.5"
services:
  orion:
    image: fiware/orion-ld
    hostname: orion
    container_name: fiware-orion
    expose:
      - "1026"
    ports:
      - "10002:1026"
    depends_on:
      - mongo-db
    command: -dbhost mongo-db -logLevel DEBUG

  mongo-db:
    image: mongo:3.6
    hostname: mongo-db
    container_name: db-mongo
    ports:
      - "27017"
    networks:
      - default
    command: --nojournal
    volumes:
      - mongo-db:/data

volumes:
  mongo-db: ~
```

Listing 2. Orion-LD docker container configuration

The following entity present in Listing 3, has been used for the development of this prototype, this entity is a sample entity given by Smart Data Models, program led by FIWARE, IUDX, TM Forum, OASC and others to support the adoption of common compatible data models in smart solutions, and it can be found in its GitHub webpage repository.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

```
{
  "id": "urn:ngsi-ld:EnvironmentObserved:{id}",
  "type": "EnvironmentObserved",
  "airQualityObserved": [
    "urn:ngsi-ld:AirQualityObserved:4b8b09c9-ce54-46de-8067-
5591e02d8f29",
    "urn:ngsi-ld:WeatherObserved:08a14933-b44d-4297-b2d2-2c3f3844012e"
  ],
  "dataProvider": "https://provider.example.com",
  "location": {
    "type": "Point",
    "coordinates": [
      -104.99404,
      39.75621
    ]
  },
  "pointOfInterest": [
    "urn:ngsi-ld:POI:cdfd9cb8-ae2b-47cb-a43a-b9767ffd5c84",
    "urn:ngsi-ld:POI:42dcd5ea-46db-11e8-bea0-772aba733f93",
    "urn:ngsi-ld:POI:4912d78e-46db-11e8-8572-ab2b8e55590b"
  ],
  "source": "https://source.example.com",
  "waterQualityObserved": [
    "urn:ngsi-ld:WeatherObserved:68a83e68-61e6-4e3c-975c-5b301c184ca6",
    "urn:ngsi-ld:WeatherObserved:b01518e3-2b60-4bbd-9783-3af0d660349e"
  ],
  "weatherObserved": [
    "urn:ngsi-ld:WeatherObserved:fae29f4c-0691-4bab-bef8-ad1cd165cc28",
    "urn:ngsi-ld:WeatherObserved:1c7a2711-ae38-4ea9-8f9f-627067067d53"
  ],
  "@context": [
    "https://raw.githubusercontent.com/smart-data-
models/dataModel.Environment/master/context.jsonld"
  ]
}
```

Listing 3. Sample entity structure for the prototype

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

The execution of the document shown in Listing 2 outputs the following messages shown in Figure 20 indicating the proper initialization of the context broker. After this output, the context broker is ready to receive requests.

```
fiware-orion  | time=Monday 28 Nov 10:35:49 2022.220Z | lvl=INFO | corr=N/A | trans=N/A | from=N/A | srv=N/A | subsrv=N/A | comp=Orion | op=orionld.cpp[1188]:main | msg=Startup completed
fiware-orion  | time=Monday 28 Nov 10:35:49 2022.221Z | lvl=TMP | corr=N/A | trans=N/A | from=N/A | srv=N/A | subsrv=N/A | comp=Orion | op=orionld.cpp[1203]:main | msg=Initialization is Done
fiware-orion  | time=Monday 28 Nov 10:35:49 2022.221Z | lvl=TMP | corr=N/A | trans=N/A | from=N/A | srv=N/A | subsrv=N/A | comp=Orion | op=orionld.cpp[1204]:main | msg=  Accepting REST requests on port 1026 (experimental API endpoints are disabled)
fiware-orion  | time=Monday 28 Nov 10:35:49 2022.221Z | lvl=TMP | corr=N/A | trans=N/A | from=N/A | srv=N/A | subsrv=N/A | comp=Orion | op=orionld.cpp[1205]:main | msg=  TRoE:                       Disabled
fiware-orion  | time=Monday 28 Nov 10:35:49 2022.221Z | lvl=TMP | corr=N/A | trans=N/A | from=N/A | srv=N/A | subsrv=N/A | comp=Orion | op=orionld.cpp[1206]:main | msg=  Forwarding:                 Disabled
fiware-orion  | time=Monday 28 Nov 10:35:49 2022.221Z | lvl=TMP | corr=N/A | trans=N/A | from=N/A | srv=N/A | subsrv=N/A | comp=Orion | op=orionld.cpp[1207]:main | msg=  Health Check:               Disabled
fiware-orion  | time=Monday 28 Nov 10:35:49 2022.221Z | lvl=TMP | corr=N/A | trans=N/A | from=N/A | srv=N/A | subsrv=N/A | comp=Orion | op=orionld.cpp[1212]:main | msg=  Mongo Server Version:       3.6.23
fiware-orion  | time=Monday 28 Nov 10:35:49 2022.221Z | lvl=TMP | corr=N/A | trans=N/A | from=N/A | srv=N/A | subsrv=N/A | comp=Orion | op=orionld.cpp[1226]:main | msg=  Mongo Driver:               Legacy C++ Driver (deprecated by mongodb)
```

Figure 20 Semantic Context Broker deployment

To test the basic functioning of the context broker, the operation shown in Section 4.2.4.1.1 will be used, usign as a payload an entity as shown in Listing 3. As it can be seen in Figure 21, the context broker outputs every entity created, as well as the code 201 indicating the proper finalization of the creation.

```
Sending to Orion-LD -> [{"id": "urn:ngsi-ld:EnvironmentObserved:33f02632-74f4-4c96-9ba1-e26945de94810", "type": "EnvironmentObserved", "airQualityObserved": ["urn:ngsi-ld:AirQualityObserved:4b8b09c9-ce54-46de-8067-5591e02d8f
29", "urn:ngsi-ld:WeatherObserved:08a14933-b44d-4297-b2d2-2c3f3844012e"], "dataProvider": "https://provider.example.com", "location": {"type": "Point", "coordinates": [-104.99404, 39.75621]}, "pointOfInterest": ["urn:ngsi-ld
:POI:cdfd9cb8-ae2b-47cb-a43a-b9767ffd5c84", "urn:ngsi-ld:POI:42dcd5ea-46db-11e8-bea0-772aba733f93", "urn:ngsi-ld:POI:4912d78e-46db-11e8-8572-ab2b8e55590b"], "source": "https://source.example.com", "waterQualityObserved": ["u
rn:ngsi-ld:WeatherObserved:68a83e68-61e6-4e3c-975c-5b301c184ca6", "urn:ngsi-ld:WeatherObserved:b01518e3-2b60-4bbd-9783-3af0d660349e"], "weatherObserved": ["urn:ngsi-ld:WeatherObserved:fae29f4c-0691-4bab-bef8-ad1cd165cc28", "
urn:ngsi-ld:WeatherObserved:1c7a2711-ae38-4ea9-8f9f-627067067d53"], "@context": ["https://raw.githubusercontent.com/smart-data-models/dataModel.Environment/master/context.jsonld"]}]

Response from Orion-LD -> code:201 message:["urn:ngsi-ld:EnvironmentObserved:33f02632-74f4-4c96-9ba1-e26945de94810"]

Sending to Orion-LD -> [{"id": "urn:ngsi-ld:EnvironmentObserved:33f02632-74f4-4c96-9ba1-e26945de94811", "type": "EnvironmentObserved", "airQualityObserved": ["urn:ngsi-ld:AirQualityObserved:4b8b09c9-ce54-46de-8067-5591e02d8f
29", "urn:ngsi-ld:WeatherObserved:08a14933-b44d-4297-b2d2-2c3f3844012e"], "dataProvider": "https://provider.example.com", "location": {"type": "Point", "coordinates": [-104.99404, 39.75621]}, "pointOfInterest": ["urn:ngsi-ld
:POI:cdfd9cb8-ae2b-47cb-a43a-b9767ffd5c84", "urn:ngsi-ld:POI:42dcd5ea-46db-11e8-bea0-772aba733f93", "urn:ngsi-ld:POI:4912d78e-46db-11e8-8572-ab2b8e55590b"], "source": "https://source.example.com", "waterQualityObserved": ["u
rn:ngsi-ld:WeatherObserved:68a83e68-61e6-4e3c-975c-5b301c184ca6", "urn:ngsi-ld:WeatherObserved:b01518e3-2b60-4bbd-9783-3af0d660349e"], "weatherObserved": ["urn:ngsi-ld:WeatherObserved:fae29f4c-0691-4bab-bef8-ad1cd165cc28", "
urn:ngsi-ld:WeatherObserved:1c7a2711-ae38-4ea9-8f9f-627067067d53"], "@context": ["https://raw.githubusercontent.com/smart-data-models/dataModel.Environment/master/context.jsonld"]}]

Response from Orion-LD -> code:201 message:["urn:ngsi-ld:EnvironmentObserved:33f02632-74f4-4c96-9ba1-e26945de94811"]

Sending to Orion-LD -> [{"id": "urn:ngsi-ld:EnvironmentObserved:33f02632-74f4-4c96-9ba1-e26945de94812", "type": "EnvironmentObserved", "airQualityObserved": ["urn:ngsi-ld:AirQualityObserved:4b8b09c9-ce54-46de-8067-5591e02d8f
29", "urn:ngsi-ld:WeatherObserved:08a14933-b44d-4297-b2d2-2c3f3844012e"], "dataProvider": "https://provider.example.com", "location": {"type": "Point", "coordinates": [-104.99404, 39.75621]}, "pointOfInterest": ["urn:ngsi-ld
:POI:cdfd9cb8-ae2b-47cb-a43a-b9767ffd5c84", "urn:ngsi-ld:POI:42dcd5ea-46db-11e8-bea0-772aba733f93", "urn:ngsi-ld:POI:4912d78e-46db-11e8-8572-ab2b8e55590b"], "source": "https://source.example.com", "waterQualityObserved": ["u
rn:ngsi-ld:WeatherObserved:68a83e68-61e6-4e3c-975c-5b301c184ca6", "urn:ngsi-ld:WeatherObserved:b01518e3-2b60-4bbd-9783-3af0d660349e"], "weatherObserved": ["urn:ngsi-ld:WeatherObserved:fae29f4c-0691-4bab-bef8-ad1cd165cc28", "
urn:ngsi-ld:WeatherObserved:1c7a2711-ae38-4ea9-8f9f-627067067d53"], "@context": ["https://raw.githubusercontent.com/smart-data-models/dataModel.Environment/master/context.jsonld"]}]

Response from Orion-LD -> code:201 message:["urn:ngsi-ld:EnvironmentObserved:33f02632-74f4-4c96-9ba1-e26945de94812"]

Sending to Orion-LD -> [{"id": "urn:ngsi-ld:EnvironmentObserved:33f02632-74f4-4c96-9ba1-e26945de94813", "type": "EnvironmentObserved", "airQualityObserved": ["urn:ngsi-ld:AirQualityObserved:4b8b09c9-ce54-46de-8067-5591e02d8f
29", "urn:ngsi-ld:WeatherObserved:08a14933-b44d-4297-b2d2-2c3f3844012e"], "dataProvider": "https://provider.example.com", "location": {"type": "Point", "coordinates": [-104.99404, 39.75621]}, "pointOfInterest": ["urn:ngsi-ld
:POI:cdfd9cb8-ae2b-47cb-a43a-b9767ffd5c84", "urn:ngsi-ld:POI:42dcd5ea-46db-11e8-bea0-772aba733f93", "urn:ngsi-ld:POI:4912d78e-46db-11e8-8572-ab2b8e55590b"], "source": "https://source.example.com", "waterQualityObserved": ["u
rn:ngsi-ld:WeatherObserved:68a83e68-61e6-4e3c-975c-5b301c184ca6", "urn:ngsi-ld:WeatherObserved:b01518e3-2b60-4bbd-9783-3af0d660349e"], "weatherObserved": ["urn:ngsi-ld:WeatherObserved:fae29f4c-0691-4bab-bef8-ad1cd165cc28", "
urn:ngsi-ld:WeatherObserved:1c7a2711-ae38-4ea9-8f9f-627067067d53"], "@context": ["https://raw.githubusercontent.com/smart-data-models/dataModel.Environment/master/context.jsonld"]}]

Response from Orion-LD -> code:201 message:["urn:ngsi-ld:EnvironmentObserved:33f02632-74f4-4c96-9ba1-e26945de94813"]

Sending to Orion-LD -> [{"id": "urn:ngsi-ld:EnvironmentObserved:33f02632-74f4-4c96-9ba1-e26945de94814", "type": "EnvironmentObserved", "airQualityObserved": ["urn:ngsi-ld:AirQualityObserved:4b8b09c9-ce54-46de-8067-5591e02d8f
29", "urn:ngsi-ld:WeatherObserved:08a14933-b44d-4297-b2d2-2c3f3844012e"], "dataProvider": "https://provider.example.com", "location": {"type": "Point", "coordinates": [-104.99404, 39.75621]}, "pointOfInterest": ["urn:ngsi-ld
:POI:cdfd9cb8-ae2b-47cb-a43a-b9767ffd5c84", "urn:ngsi-ld:POI:42dcd5ea-46db-11e8-bea0-772aba733f93", "urn:ngsi-ld:POI:4912d78e-46db-11e8-8572-ab2b8e55590b"], "source": "https://source.example.com", "waterQualityObserved": ["u
rn:ngsi-ld:WeatherObserved:68a83e68-61e6-4e3c-975c-5b301c184ca6", "urn:ngsi-ld:WeatherObserved:b01518e3-2b60-4bbd-9783-3af0d660349e"], "weatherObserved": ["urn:ngsi-ld:WeatherObserved:fae29f4c-0691-4bab-bef8-ad1cd165cc28", "
urn:ngsi-ld:WeatherObserved:1c7a2711-ae38-4ea9-8f9f-627067067d53"], "@context": ["https://raw.githubusercontent.com/smart-data-models/dataModel.Environment/master/context.jsonld"]}]

Response from Orion-LD -> code:201 message:["urn:ngsi-ld:EnvironmentObserved:33f02632-74f4-4c96-9ba1-e26945de94814"]
```

Figure 21 Prototype output entities creation

This project has received funding from the European Union's Horizon 2020 research
and innovation programme under grant agreement No. 101037084.

To recover the entities previously created, the Entity Reading operation defined in Section 4.2.4.1.1 can be used to recover all the entities by its type. The output of this operation is shown in Figure 22, this output is formed by the code 200 and a list with each entity created.


```
Requesting Enties fromOrion-LD

Response from Orion-LD -> code:200 message:[{'id': 'urn:ngsi-ld:EnvironmentObserved:33f02632-74f4-4c96-9ba1-e26945de94810', 'type': 'EnvironmentObserved', 'https://smartdatamodels.org/dataModel.Environment/airQualityObserved
': {'type': 'Property', 'value': ['urn:ngsi-ld:AirQualityObserved:4b8b09c9-ce54-46de-8067-5591e02d8f29', 'urn:ngsi-ld:WeatherObserved:08a14933-b44d-4297-b2d2-2c3f3844012e']}, 'https://smartdatamodels.org/dataModel.Environmen
t/pointOfInterest': {'type': 'Property', 'value': ['urn:ngsi-ld:POI:cdfd9cb8-ae2b-47cb-a43a-b9767ffd5c84', 'urn:ngsi-ld:POI:42dcd5ea-46db-11e8-bea0-772aba733f93', 'urn:ngsi-ld:POI:4912d78e-46db-11e8-8572-ab2b8e55590b']}, 'ht
tps://smartdatamodels.org/dataModel.Environment/waterQualityObserved': {'type': 'Property', 'value': ['urn:ngsi-ld:WeatherObserved:68a83e68-61e6-4e3c-975c-5b301c184ca6', 'urn:ngsi-ld:WeatherObserved:b01518e3-2b60-4bbd-9783-3
af0d660349e']}, 'https://smartdatamodels.org/dataModel.Environment/weatherObserved': {'type': 'Property', 'value': ['urn:ngsi-ld:WeatherObserved:fae29f4c-0691-4bab-bef8-ad1cd165cc28', 'urn:ngsi-ld:WeatherObserved:1c7a2711-ae
38-4ea9-8f9f-627067067d53']}, 'https://smartdatamodels.org/dataProvider': {'type': 'Property', 'value': 'https://provider.example.com'}, 'https://smartdatamodels.org/source': {'type': 'Property', 'value': 'https://source.exa
mple.com'}, 'location': {'type': 'GeoProperty', 'value': {'type': 'Point', 'coordinates': [-104.99404, 39.75621]}}}, {'id': 'urn:ngsi-ld:EnvironmentObserved:33f02632-74f4-4c96-9ba1-e26945de94811', 'type': 'EnvironmentObserve
d', 'https://smartdatamodels.org/dataModel.Environment/airQualityObserved': {'type': 'Property', 'value': ['urn:ngsi-ld:AirQualityObserved:4b8b09c9-ce54-46de-8067-5591e02d8f29', 'urn:ngsi-ld:WeatherObserved:08a14933-b44d-429
7-b2d2-2c3f3844012e']}, 'https://smartdatamodels.org/dataModel.Environment/pointOfInterest': {'type': 'Property', 'value': ['urn:ngsi-ld:POI:cdfd9cb8-ae2b-47cb-a43a-b9767ffd5c84', 'urn:ngsi-ld:POI:42dcd5ea-46db-11e8-bea0-772
aba733f93', 'urn:ngsi-ld:POI:4912d78e-46db-11e8-8572-ab2b8e55590b']}, 'https://smartdatamodels.org/dataModel.Environment/waterQualityObserved': {'type': 'Property', 'value': ['urn:ngsi-ld:WeatherObserved:68a83e68-61e6-4e3c-9
75c-5b301c184ca6', 'urn:ngsi-ld:WeatherObserved:b01518e3-2b60-4bbd-9783-3af0d660349e']}, 'https://smartdatamodels.org/dataModel.Environment/weatherObserved': {'type': 'Property', 'value': ['urn:ngsi-ld:WeatherObserved:fae29f
4c-0691-4bab-bef8-ad1cd165cc28', 'urn:ngsi-ld:WeatherObserved:1c7a2711-ae38-4ea9-8f9f-627067067d53']}, 'https://smartdatamodels.org/dataProvider': {'type': 'Property', 'value': 'https://provider.example.com'}, 'https://smart
datamodels.org/source': {'type': 'Property', 'value': 'https://source.example.com'}, 'location': {'type': 'GeoProperty', 'value': {'type': 'Point', 'coordinates': [-104.99404, 39.75621]}}}, {'id': 'urn:ngsi-ld:EnvironmentObs
erved:33f02632-74f4-4c96-9ba1-e26945de94812', 'type': 'EnvironmentObserved', 'https://smartdatamodels.org/dataModel.Environment/airQualityObserved': {'type': 'Property', 'value': ['urn:ngsi-ld:AirQualityObserved:4b8b09c9-ce5
4-46de-8067-5591e02d8f29', 'urn:ngsi-ld:WeatherObserved:08a14933-b44d-4297-b2d2-2c3f3844012e']}, 'https://smartdatamodels.org/dataModel.Environment/pointOfInterest': {'type': 'Property', 'value': ['urn:ngsi-ld:POI:cdfd9cb8-a
e2b-47cb-a43a-b9767ffd5c84', 'urn:ngsi-ld:POI:42dcd5ea-46db-11e8-bea0-772aba733f93', 'urn:ngsi-ld:POI:4912d78e-46db-11e8-8572-ab2b8e55590b']}, 'https://smartdatamodels.org/dataModel.Environment/waterQualityObserved': {'type'
: 'Property', 'value': ['urn:ngsi-ld:WeatherObserved:68a83e68-61e6-4e3c-975c-5b301c184ca6', 'urn:ngsi-ld:WeatherObserved:b01518e3-2b60-4bbd-9783-3af0d660349e']}, 'https://smartdatamodels.org/dataModel.Environment/weatherObse
rved': {'type': 'Property', 'value': ['urn:ngsi-ld:WeatherObserved:fae29f4c-0691-4bab-bef8-ad1cd165cc28', 'urn:ngsi-ld:WeatherObserved:1c7a2711-ae38-4ea9-8f9f-627067067d53']}, 'https://smartdatamodels.org/dataProvider': {'ty
pe': 'Property', 'value': 'https://provider.example.com'}, 'https://smartdatamodels.org/source': {'type': 'Property', 'value': 'https://source.example.com'}, 'location': {'type': 'GeoProperty', 'value': {'type': 'Point', 'co
ordinates': [-104.99404, 39.75621]}}}, {'id': 'urn:ngsi-ld:EnvironmentObserved:33f02632-74f4-4c96-9ba1-e26945de94813', 'type': 'EnvironmentObserved', 'https://smartdatamodels.org/dataModel.Environment/airQualityObserved': {'
type': 'Property', 'value': ['urn:ngsi-ld:AirQualityObserved:4b8b09c9-ce54-46de-8067-5591e02d8f29', 'urn:ngsi-ld:WeatherObserved:08a14933-b44d-4297-b2d2-2c3f3844012e']}, 'https://smartdatamodels.org/dataModel.Environment/poi
ntOfInterest': {'type': 'Property', 'value': ['urn:ngsi-ld:POI:cdfd9cb8-ae2b-47cb-a43a-b9767ffd5c84', 'urn:ngsi-ld:POI:42dcd5ea-46db-11e8-bea0-772aba733f93', 'urn:ngsi-ld:POI:4912d78e-46db-11e8-8572-ab2b8e55590b']}, 'https:/
/smartdatamodels.org/dataModel.Environment/waterQualityObserved': {'type': 'Property', 'value': ['urn:ngsi-ld:WeatherObserved:68a83e68-61e6-4e3c-975c-5b301c184ca6', 'urn:ngsi-ld:WeatherObserved:b01518e3-2b60-4bbd-9783-3af0d6
60349e']}, 'https://smartdatamodels.org/dataModel.Environment/weatherObserved': {'type': 'Property', 'value': ['urn:ngsi-ld:WeatherObserved:fae29f4c-0691-4bab-bef8-ad1cd165cc28', 'urn:ngsi-ld:WeatherObserved:1c7a2711-ae38-4e
a9-8f9f-627067067d53']}, 'https://smartdatamodels.org/dataProvider': {'type': 'Property', 'value': 'https://provider.example.com'}, 'https://smartdatamodels.org/source': {'type': 'Property', 'value': 'https://source.example.
com'}, 'location': {'type': 'GeoProperty', 'value': {'type': 'Point', 'coordinates': [-104.99404, 39.75621]}}}, {'id': 'urn:ngsi-ld:EnvironmentObserved:33f02632-74f4-4c96-9ba1-e26945de94814', 'type': 'EnvironmentObserved', '
https://smartdatamodels.org/dataModel.Environment/airQualityObserved': {'type': 'Property', 'value': ['urn:ngsi-ld:AirQualityObserved:4b8b09c9-ce54-46de-8067-5591e02d8f29', 'urn:ngsi-ld:WeatherObserved:08a14933-b44d-4297-b2d
2-2c3f3844012e']}, 'https://smartdatamodels.org/dataModel.Environment/pointOfInterest': {'type': 'Property', 'value': ['urn:ngsi-ld:POI:cdfd9cb8-ae2b-47cb-a43a-b9767ffd5c84', 'urn:ngsi-ld:POI:42dcd5ea-46db-11e8-bea0-772aba73
3f93', 'urn:ngsi-ld:POI:4912d78e-46db-11e8-8572-ab2b8e55590b']}, 'https://smartdatamodels.org/dataModel.Environment/waterQualityObserved': {'type': 'Property', 'value': ['urn:ngsi-ld:WeatherObserved:68a83e68-61e6-4e3c-975c-5
b301c184ca6', 'urn:ngsi-ld:WeatherObserved:b01518e3-2b60-4bbd-9783-3af0d660349e']}, 'https://smartdatamodels.org/dataModel.Environment/weatherObserved': {'type': 'Property', 'value': ['urn:ngsi-ld:WeatherObserved:fae29f4c-06
91-4bab-bef8-ad1cd165cc28', 'urn:ngsi-ld:WeatherObserved:1c7a2711-ae38-4ea9-8f9f-627067067d53']}, 'https://smartdatamodels.org/dataProvider': {'type': 'Property', 'value': 'https://provider.example.com'}, 'https://smartdatam
odels.org/source': {'type': 'Property', 'value': 'https://source.example.com'}, 'location': {'type': 'GeoProperty', 'value': {'type': 'Point', 'coordinates': [-104.99404, 39.75621]}}}]
```


Figure 22 Prototype output entities reading

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

6 Conclusions & future work

The final chapter of the deliverable is mainly focused on the description of the different conclusions and results obtaining during the execution of the Task 2.2 entitled *“Open data space for enabling knowledge sharing”* presented in Section 2.1. This task started on M8 (May-2022) and is expected to finish on M36 (30 September 2023). A second version of this deliverable is also expected, when all the results and outcomes from tasks presented in Section 6.2 will be acquired. Also, this second version of the document will focus on the presentation of the constructed knowledge graph, as well as the refinement of the context broker in relation of the requirements of the architecture.

6.1 Conclusions

The present deliverable has provided a wide vision about the **semantic context broker** to be adopted in IMPETUS project. As depicted at the beginning of the document, this tool has the main objective of **establishing the technological bases for data integration from case-studies combined with the integration of data from public and private repositories**.

For that, the semantic context broker is a continuation of the work presented in D2.1. In that document, it was established the requirements and needs for the resilient knowledge boosters, including an architecture of the whole platform. One of the main pieces of that platform is the semantic-context broker that will enable to create **an open data space for climate** in the different biogeographical regions included in IMPETUS.

To elaborate such tools, the document reviewed the technological stack and approaches in relation to this trending topic that is arising in Europe. The main conclusion about the revision analysis is that the **open data spaces are still a new concept that is continuously maturing**. However, IDSA and FIWARE approaches are the prominent technology to become the standard de facto for the data spaces concept. For that reason, one of our **decisions is to make the semantic context broker compatible with FIWARE technology**.

After taking this decision, the rest of the document presented our approach to integrate climate related data. The data repositories to be integrated is a challenge due to their sizes ranging from 10Gb to 1TB of data. Therefore, we need to elaborate a context broker that goes beyond the state of the art in the following points:

- **Reduce time consuming in data integration or consumption** when huge amounts of information are integrated into the context broker.
- **Full semantic reasoning** inside knowledge graphs to understand information linkage and potential simple causal events in climate using inductive and probabilistic reasoning.
- **Stream processing** coming from repositories, or the joining integration actions from the IoT Agents.
- **Enabling semantic rules** for facilitating mappings coming from information in different types (enabling R2RML rules for data integration)
- **Enabling rules for data validation** using, for example, SHACL rules.

The constructed context broker has been initially tested but still requires some future advancements and refinement when we include public and private data repositories.

6.2 Future work

T2.2 will continue developments until M30 (March-2024). At this deadline, we are expecting the presentation of D2.7 Semantic Context Broker tool v2, where all the evolution of the context broker as well as the development of the ontology and the semantic repository will be reflected. The following table (Table 17) presents the expected development to be presented in D2.7:

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 101037084.

Item	Description
Semantic Context Broker	- Development of an strategy to deploy instances from the context brokers. - Configuration and deployment of the context broker for the different biogeographical regions - Private and public data integration - Implementation of Owner and access to the different repositories - Incorporation of different semantic repository databases (Quad stores). - Incorporation of needed mechanism to enable queries with reasoning. - Incorporation of geospatial reasoning and querying about the information - Incorporation of time-databases to store long-term time series of historical information.
Ontology	- Publication of the initial version of the ontology - Incorporation of SHACL rules for validating data and the data models - Incorporation of SWRL rules to provide richness to the reasoning.
IoT Agents	- Incorporation of R2RML rules to transform information automatically with the corresponding context.

Table 17 Future work to be performed in relation to IMPETUS context broker, ontology, and semantic repository

At IMPETUS general level, the advancements and deployment of the context-broker will establish the basis for the data collection and manipulations. Thus, the context-broker will suppose the base for the elaboration of the AI driven tools and visualization engines. So, the specific contribution of the context-broker will impact on:

Table 18. Impact on context broker advancements in WPs

Workpackage	Description
WP2	- Support in the elaboration of the AI driven tools - Support in the development of the visualization engine.
WP3	- Climatic scenarios and projections data integration - Data integration for risk assessment
WP4	- Data presentation to the demo-cases (regionaities) - Data integration at regional scale
WP5	- Bring a tool to support the elaboration of the adaptation strategies and innovation packages.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101037084.

7 References

(UiT), A. M. (2022). *Design strategy of the digital dimension of Resilience Knowledge Boosters.*

ETSI, G. (2019). *Context Information Management (CIM); NGSI-LD API.* Tech. Rep.[Online]. Available: https://www.etsi.org/deliver/etsi_gs/CIM/001_099/009/01.01.01_60/gs_cim009v010101p.pdf.

European Commission, Directorate-General for Communications Networks, Content and Technology. (2018). *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Artificial Intelligence for Europe.*

European Commission, Directorate-General for Communications Networks, Content and Technology. (2020). *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS A European strategy for data.*

European Commission, Directorate-General for Communications Networks, Content and Technology. (2020). *Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on European data governance (Data Governance Act).*

European Parliament, Council of the European Union. (2019). *Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (recast).*

Edward Curry, S. S. (2022). *Data Spaces: Design, Deployment, and Future Directions.* Springer.

Edward Curry, S. S. (2022). Data Spaces Design, Deployment and Future Directions. Springer.

United Nations. (2005, 26 03). *unece.* Retrieved 12 15, 2022, from https://www.unece.org/fileadmin/DAM/cefact/recommendations/rec20/rec20_rev3_Annex3e.pdf

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 101037084.