FIWARE FOR DATA SPACES

POSITION PAPER / VERSION 1.0 / JUNE 2021

FIWARE
Open APIs for Open Minds

www.fiware.org

Publisher

FIWARE Foundation e.V.
Franklinstraße 13A
10587 Berlin
Germany

Editor

Juanjo Hierro
FIWARE Foundation

Copyright

Disclaimer

In accordance with our Guidelines concerning the use of technology, strategic and market information in positioning and advertising, please be aware of the following: Position Papers appearing on the FIWARE Foundation site or in other digital or printed materials are actually received via text, audio or video submission. They are individual experiences and common strategies, reflecting studies and real life experiences of those who will use or have used our technology and/or services in some way or another. We do not claim that they are typical results that customers will generally use and achieve. Some FIWARE Position Papers have been shortened.

Table of contents

Overview

This paper describes how smart applications from multiple domains can participate in the creation of Data Spaces based on FIWARE software Building Blocks. Smart applications participating in such Data Spaces share Digital Twin data in real-time using a common standard API like NGSI-LD and relying on standard data models. Each smart solution contributes to build a complete Digital Twin data representation of the real world sharing their data. At the same time, they can exploit data shared by other applications. Relying on FIWARE Data Marketplace components, smart applications can publish data under concrete terms and conditions which include pricing or data usage/access policies.

A federated cloud infrastructure and mechanisms supporting data sovereignty and trust are necessary to create Data Spaces. However, additional elements have to be added to ease the creation of data value chains and the materialization of a data economy. Standard APIs, combined with standard data models, are crucial to support effective data exchange enabling loose coupling between parties as well as reusability and replaceability of data resources and applications. Similarly, Data Spaces need to incorporate mechanisms for publication, discovery and trading of data resources. These are elements that FIWARE implements and can be combined with architecture elements like the IDS Connector proposed by IDSA or iSHARE Satellite technology to create Data Spaces supporting trusted and effective data sharing.

The alignment of FIWARE Building Blocks for Data Spaces with Connecting Europe Facility Building Blocks is essential when considering creation of Data Spaces in Europe. FIWARE mature technologies, integrated with IDS Connector technologies can contribute to accelerate materialization of GAIA-X, a project started in 2020 aimed at creating a federated form of data infrastructure in Europe that strengthens the ability to both access and share data securely and confidently.

1 Introduction to FIWARE

FIWARE was created with the ultimate goal of creating an open sustainable ecosystem around public, royalty-free and implementation-driven software platform standards easing the development of smart solutions and supporting organizations in their transition into smart organizations. From a technical perspective, FIWARE brings a curated framework of Open Source software components which can be assembled together and combined with other third-party platform components to build platforms easing the development of smart solutions and smart organizations in multiple application domains: cities, manufacturing, utilities, agrifood, etc. Since creation of the FIWARE Foundation in late 2016, a vibrant FIWARE Community has been formed with a true worldwide dimension, comprising more than 90 member organizations, including large corporations, SMEs, technology centres and universities, and hundreds of individual members. Parallel to this growth the number of organizations adopting FIWARE has not stopped increasing.

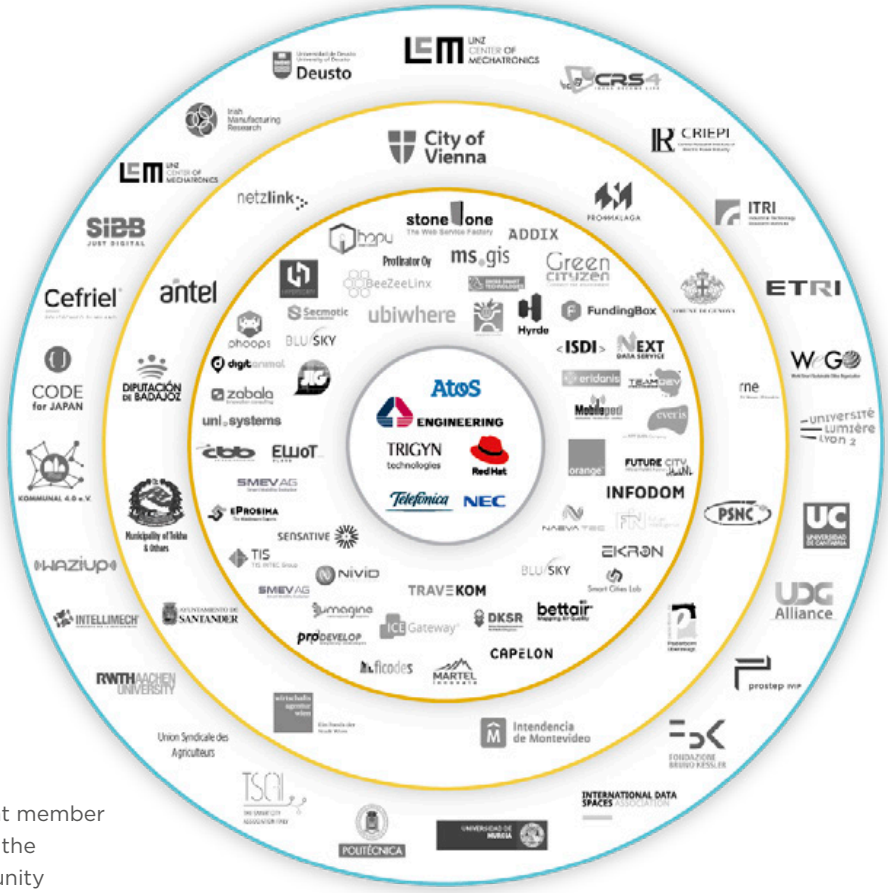

Figure 1 - Current member organizations in the FIWARE Community

Any software architecture "powered by FIWARE" gravitates around management of a Digital Twin data representation of the real world. This Digital Twin data representation is built based on information gathered from many different sources, including sensors, cameras, information systems, social networks, end users through mobile devices, etc. It is constantly maintained and accessible in near real-time ("right-time" is the term also often used, reflecting that the interval between the instants of time at which some data is gathered and made accessible is enough short to allow a proper reaction). Applications constantly process and analyze this data (not only current values but also history generated over time) in order to automate certain tasks or bring support to smart decisions by end users. The collection of all Digital Twins modelling the real world that is managed is also referred to as **Context** and the data associated with attributes of Digital Twins is also referred to as **context information**.

In FIWARE, a Digital Twin is an entity which digitally represents a real-world physical asset (e.g. a bus in a city, a milling machine in a factory) or a concept (e.g., a weather forecast, a product order). Each Digital Twin:

- is universally identified with a URI (Universal Resource Identifier),
- belongs to a well-known type (e.g., the Bus type, of the Room type) also universally identified by a URI, and
- is characterized by several attributes which in turn are classified as:
  - properties holding data (e.g., the "current speed" of a Bus, or "max temperature" in a Room) or
  - relationships, each holding a URI identifying another Digital Twin entity the given entity is linked to (e.g., the concrete Building where a concrete Room is located).

Attributes of a Digital Twin may vary ranging from attributes that are quite static (e.g., the "license plate" of a Bus), to attributes that change very dynamically (e.g., the "speed" or "number of passengers" in a Bus) to attributes which still change but not that often (e.g., the "driver" in a Bus which may change twice a day). Very important, the attributes of a Digital Twin are not only limited to observable data but also inferred data. Thus, for example, the Digital Twin of a Street may not only have an attribute "current traffic", which may be automatically measured through sensors or cameras, but an attribute "forecasted traffic in 30 minutes" which might be calculated based on AI algorithms that keep the value of this attribute updated based on current traffic data, other relevant data impacting traffic (e.g., current weather observed and forecasted, schedule of events nearby, etc) and historical information about traffic in the given street. Therefore, the Digital Twin data representation of the world that is managed in a "powered by FIWARE" architecture is expected to contain all the information needed by smart applications, not only measurable data but also other augmented insights and knowledge acquired over time.

A Digital Twin approach provides the basis for data integration at different levels, as illustrated in *Figure 2*:

- Within a **vertical Smart Solution**, solving how main Building Blocks within the architecture can be integrated;.
- Within a **Smart Organization**, bringing support to the integration of the different sys-

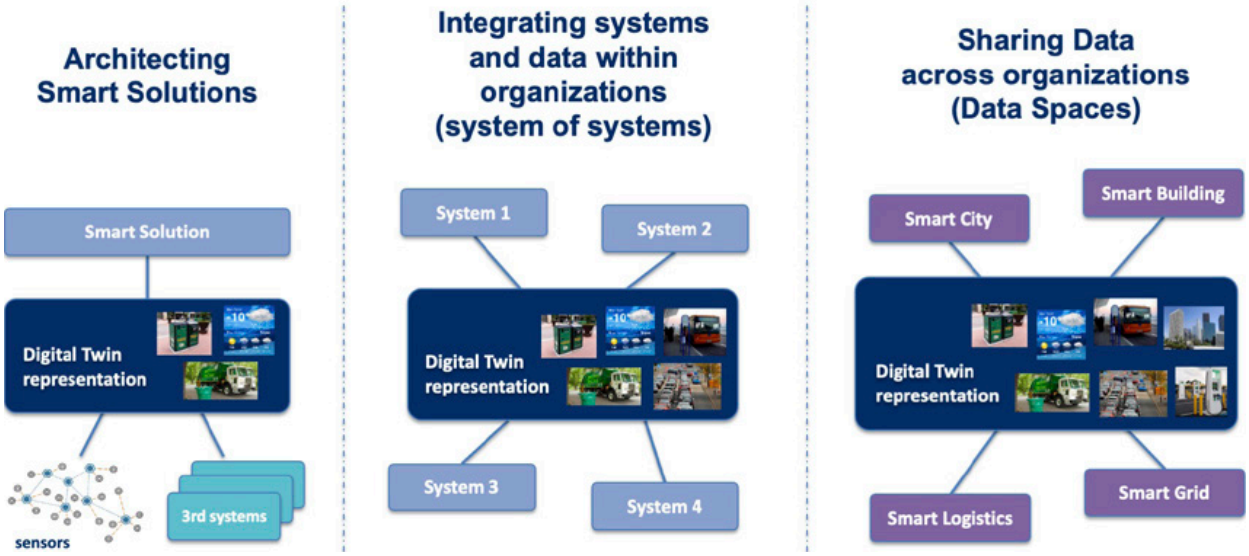

Figure 2 - Levels of Integration supported following a Digital Twin approach

tems within a smart organization following a system of systems approach;

- Within a **Smart Data Space**, establishing the basic "common lingua" that systems linked to the different organizations speak and understand.

Two critical elements need to be standardized in order to support an effective data integration at these three levels: the API to get access to Digital Twin data and the data models describing the attributes and semantics associated with the different types of Digital Twins being considered. The FIWARE Community has driven and continues to drive standardization at both fronts:

- The **NGSI API** provides a simple yet powerful RESTful API for getting access to context / Digital Twin data. NGSI API specifications have evolved over time driven by feedback from developers and implementation experiences. A first mature version of the API was the **NGSIv2 API**, which was defined by members of the FIWARE Community and is currently used in many systems in production within multiple sectors. Evolution of the API has taken place within the ETSI ISG CIM (Context Information Management Industry Specification Group), where members of the FIWARE Community and the FIWARE Foundation have led the definition of an evolved version of the API, known as the NGSI-LD API, whose specifications were first published by ETSI in 2019 and continue to evolve. The NGSI-LD API is used as the data integration API and is implemented by the core component of any "powered by FIWARE" architecture: the so-called Context Broker component. Different alternative Open Source implementations of a Context Broker are available within the FIWARE Community, namely the Orion-LD, Scorpio and Stellio products.
- The **Smart Data Models initiative** (website, github), launched by the FIWARE Foundation, provides a library of Data Models described in

JSON/JSON-LD format which are compatible respectively with the NGSIv2/NGSI-LD APIs or would be useful for defining other RESTful interfaces for accessing Digital Twin data. Data models published under the initiative are compatible with schema.org and comply with other existing de-facto sectoral standards when they exist. They solve one major issue developers are facing, that is the fact that a given data model specification may be mapped into JSON/JSON-LD in many different ways, all of them valid. Thanks to the Smart Data Models initiative, developers can rely on concrete mappings into JSON/JSON-LD, compatible with the NGSIv2/NGSI-LD APIs, that are made available within this library, avoiding interoperability problems derived from alternative mappings. Since its creation, more than 500 data models have been published and the number of organizations contributing data model descriptions is constantly growing. Relevant organizations like TM Forum, OASC or IUDX are joining forces with the FIWARE Foundation bringing support to an open governance model for the initiative, following best Open Source practices.

Building around the FIWARE Context Broker, which is the core mandatory component in any "powered by FIWARE" architecture, a rich set of complementary open-source components are available listed as part of the FIWARE Catalogue. These components can be classified in the following categories or chapters:

- components easing development of interfaces with the Internet of Things, Robotic and third-party systems;
- components supporting context / Digital Twin data processing and monitoring or the connection with data processing engines (e.g., Apache Spark, Apache Flink), monitoring tools (e.g., Grafana) and analysis platforms (e.g., Apache Superset);
- components covering aspects related to Identity and Access Management (IAM) as well as Publication and Monetization of Data (including data accessible via APIs like NGSI-LD).

Figure 3 depicts the reference architecture of a vertical smart solution powered by FIWARE. The concrete example corresponds to a smart solution for picking and palletizing products from a warehouse using robots. This reference architecture is structured in essentially three layers:

- A Context Broker component is at the core of the architecture, keeping a Digital Twin representation of the real world objects and concepts relevant to the specific problem tackled: AGV robots, palletizer robots, shelf sections where products are stored in the warehouse, automatic doors AGV robots have to pass, operators in the shopfloor, items of stored products, orders generated from the CRM system, etc.
- Southbound to the Context Broker, the NGSI IoT Agents, available as part of the FIWARE IDAS framework, are used for connections to robotic systems exporting the OPC-UA IoT protocol or to specific sensors or actuators, used for example to detect items in shelf sections or to be able to open the shop floor doors. They perform the necessary conversions between IoT protocols and NGSI. In addition, System Adapters developed based on the IDAS Agent library cope with the connection to the CRM and the Warehouse Inventory Management system that the solution has to interface with. FIWARE components like

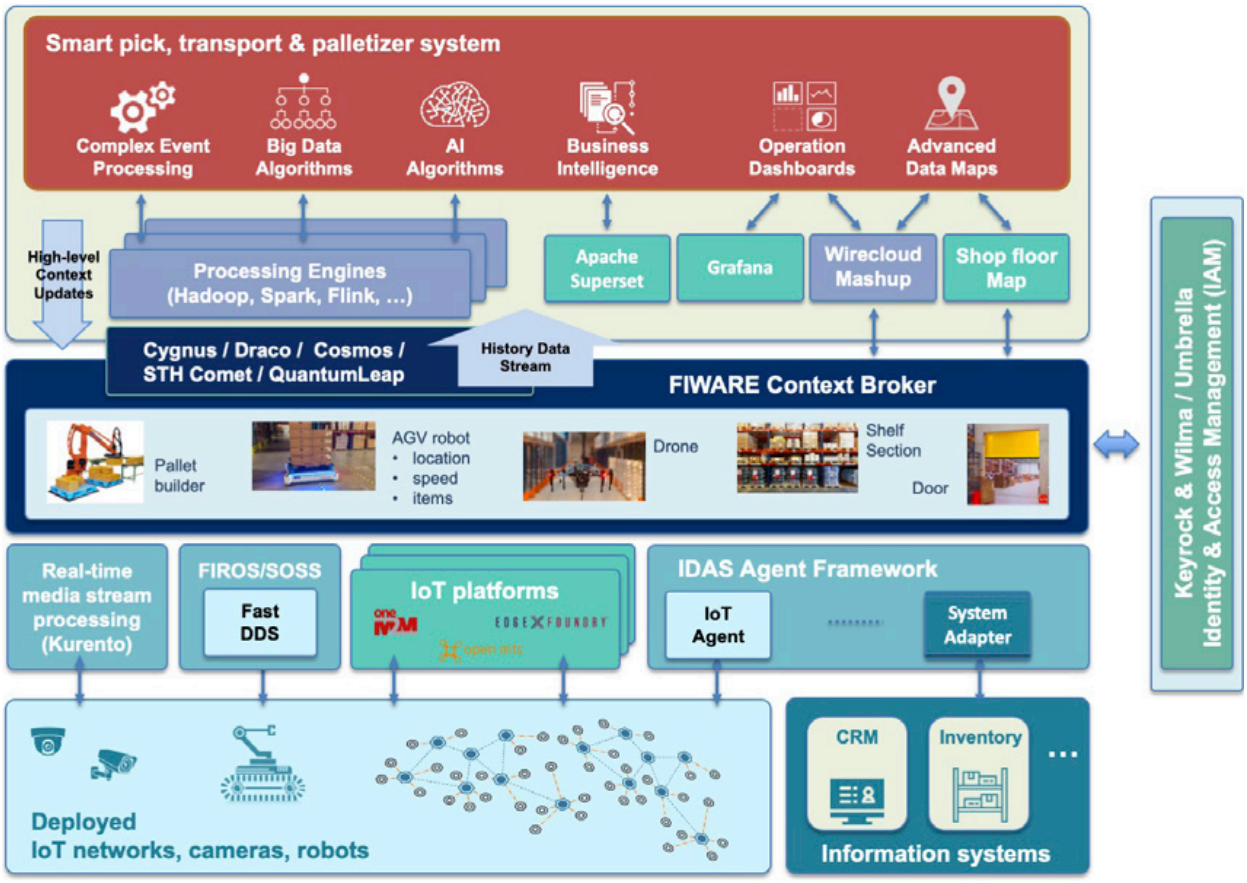

Figure 3 - Smart Solution for Picking and Palletizing products from a warehouse

FIROS/SOSS allow, on the other hand, to perform the adaptation to robotic systems based on ROS/ROS2. Last but not least, the FIWARE component Kurento is able to process the video streams of cameras deployed in the shop floor, which are helpful to detect potential obstacles or risky situations.

- Northbound to the Context Broker, a number of tools are targeted to support real-time big data processing of the streams of history data generated as context / Digital Twin information evolves over time. A combination of Open Source components from third party products (Apache Superset, Grafana) and FIWARE (e.g., Wirecloud) are shown in the picture targeted to support the creation of operational dashboards and advanced data maps for monitoring processes. A number of FIWARE Data Connectors (Cygnus, Draco, Cosmos, STH Comet, QuantumLeap) are available as part of FIWARE to facilitate transference of historic context / Digital Twin information to these tools.

Transversal to all these layers, a number of FIWARE components support Identity and Access Management (e.g., Keyrock, API Umbrella, AuthZForce). They control the flow of data across the different layers. With regards to the

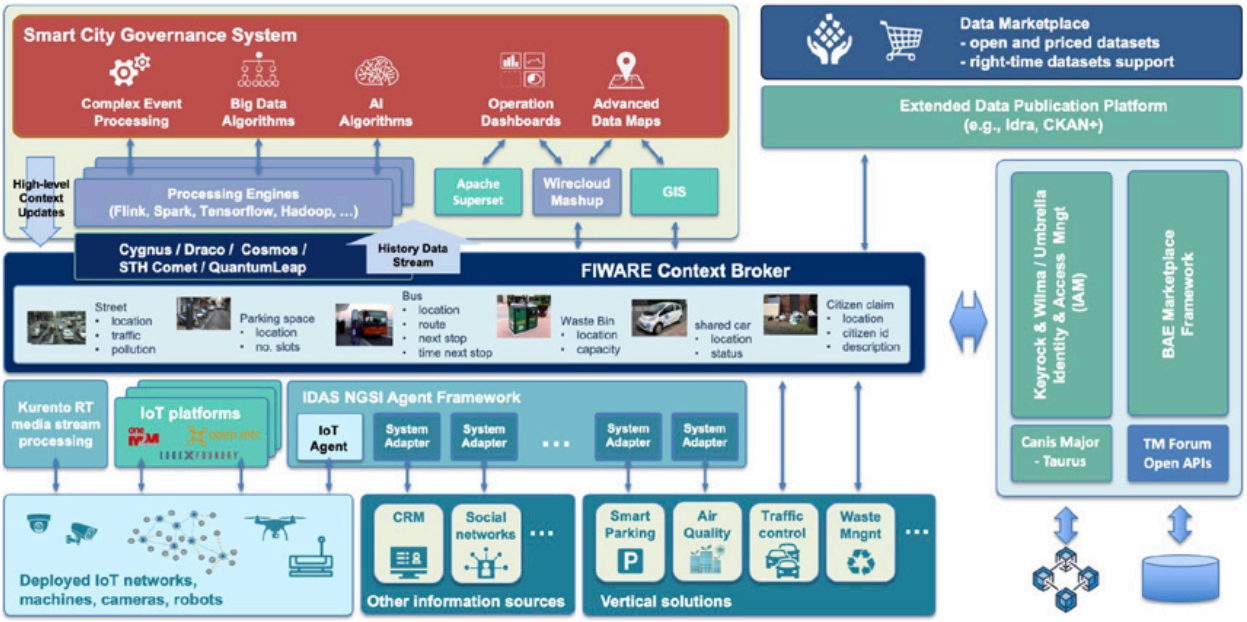

Figure 4 - Smart City Reference Architecture

access to the Context Broker, they enforce the policies establishing what users can update, query or subscribe to changes on context / Digital Twin data. Note that the flow of data is not only south to north in the picture. Northbound applications can perform updates on context data, which in turn will trigger changes in the devices, robots or systems that are connected southbound.

An important point to highlight is that FIWARE is not about taking it all or nothing. You are not forced to use all the complementary FIWARE components mentioned above but you are free to use other third party platform components as well to design the hybrid platform of your choice. Thus, for example, you may opt for using a concrete IoT platform instead of IDAS IoT Agents to interface with sensors and actuators as reflected in the picture. As long as it uses the FIWARE Context Broker technology to manage context information, your platform can be labeled as "powered by FIWARE" and solutions built on top as well.

Figure 4 depicts the reference architecture of a smart city powered by FIWARE. Again, the Context Broker component is at the core of the architecture, holding a Digital Twin representation of the real world objects and concepts and describing what is going on in the city: streets, waste bins and containers, waste trucks, buses, electric vehicle chargers, buildings, events, citizen claims, etc. The different vertical smart solutions deployed in the city (e.g., Air Quality Monitoring, Smart Traffic Management, Smart Parking, Smart Waste Management) are connected to the Context Broker contributing that information they manage which is relevant for creating a holistic Context / Digital

Twin representation of the whole City, thereby breaking the information silos. Some of these vertical smart solutions may be powered by FIWARE (e.g., Traffic Control and Waste Management systems in the figure) in which case their interface with the global city-level Context Broker does not require any adaptation. Others may not be powered by FIWARE but this doesn't represent a major problem because creation of NGSI system adapters which translate from whatever API those systems export to NGSI-LD has proven not to be difficult. Last but not least, the City may deploy sensor/camera infrastructures through which valuable data is extracted.

Exploiting the complete Context / Digital Twin representation of the City, the Smart City Governance System (or City Operation Center) can be developed. Real-time BigData processing tools can be used relying on data coming from multiple sources, extracting more valuable insights for the support of decisions. Similarly, monitoring tools can leverage this holistic Context / Digital Twin representation of the City.

So far, we have described in more detail how FIWARE components can be used to achieve the two first levels of data integration using a Digital Twin approach (see Figure 2). Within the next section, we will elaborate on how FIWARE also helps to achieve the third level of data integration, linked to the creation of Data Spaces as a natural extension of the first two. Actually, in figure 4, part of the Context / Digital Twin data representation of the city can be made available to Data Spaces, published through Data Marketplaces.

2 FIWARE and Data Spaces

A **Data Space** can be defined as a decentralized data ecosystem built around commonly agreed Building Blocks enabling an effective and trusted sharing of data among participants. From a technical perspective, a number of **technology Building Blocks** are required ensuring:

- **Data interoperability.** Data Spaces should provide a solid framework for an efficient exchange of data among participants, supporting full decoupling of data providers and consumers. This requires the adoption of a "common lingua" every participant uses, materialized in the adoption of common APIs for the data exchange, and the definition of common data models. Common mechanisms for traceability of data exchange transactions and data provenance, are also required.
- **Data sovereignty and trust.** Data Spaces should bring technical means for guaranteeing that participants in a Data Space can trust each other and exercise sovereignty over data they share. This requires the adoption of common standards for managing the identity of participants, the verification of their truthfulness and the enforcement of policies agreed upon data access and usage control.
- **Data value creation.** Data Spaces should provide support for the creation of multi-sided markets where participants can generate value out of sharing data (i.e., creating data value chains). This requires the adoption of common mechanisms enabling the definition of terms and conditions (including pricing) linked to data offerings, the publication and discovery of such offerings and the management of all the necessary steps supporting the lifecycle of contracts that are established when a given participant acquires the rights to access and use data.

Besides the adoption of a common technology foundation, Data Spaces also require **governance**, that is the adoption of a number of business, operational and organizational agreements among participants. Business agreements, for example, specify what kind of terms and conditions can regulate the sharing of data between participants and the legal framework supporting contracts established through the Data Space. Operational agreements, on the other hand, regulate policies that have to be enforced during Data Space operation like, for example, compliance with GDPR (General Data Protection Regulation) or the 2nd Payment Services Directive (PSD2) in the finance sector. They may also comprise the definition of tools that operators of cloud infrastructures or global services supporting Data Spaces must implement, enabling auditing of certain processes or the adoption of cyber-security practices. Last but not least, organizational agreements establish the governance bodies (very much like ICANN for the Internet). They deal with the identification of concrete specifications that products implementing technology Building Blocks in a Data Space should comply with, as well as the business and operational agreements to be adopted. The complete taxonomy of Building Blocks required for creating Data Spaces is illustrated in *Figure 5*. This same taxonomy of building blocks is also described in the Open DEI White Paper on Design Principles for Data Spaces.

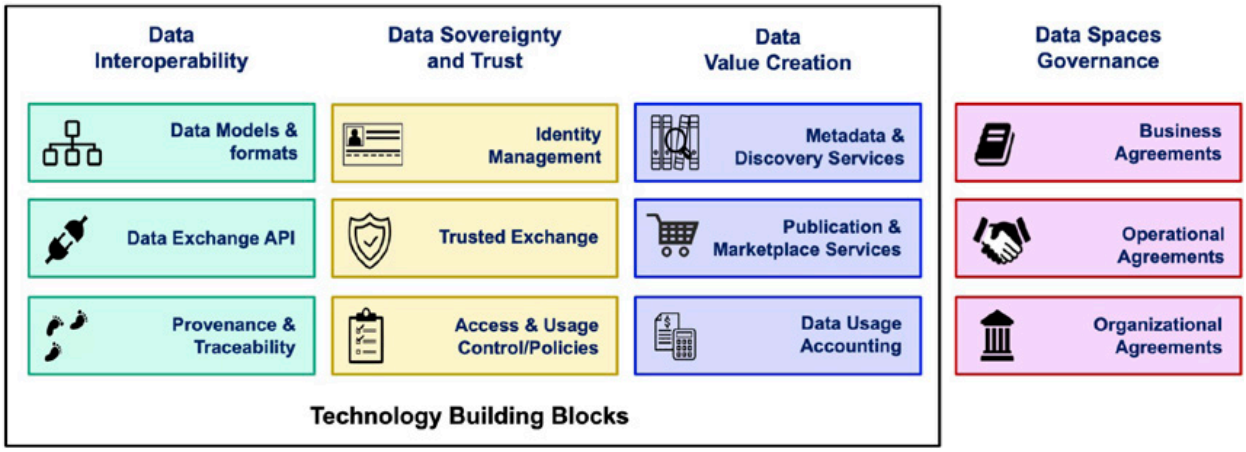

Figure 5 - Building Blocks in a Data Space

Sharing of data within a given Data Space should not be limited to a single domain. This would severely limit the creation of new innovative services since individuals and organizations usually act in multiple domains at the same time and many opportunities will flourish when data generated within organizations operating in certain domain (management of traffic in cities, for example) is shared for its exploitation in processes relevant to other domains (continuing with the example, logistics). Therefore, technology Building Blocks for Data Spaces must be domain-agnostic. On the other hand, they should rely on open standards, allowing multiple infrastructure and global service providers to emerge and support Data Spaces, without getting locked by any particular provider. Given this, while making things work in living labs and pilots is relatively easy, the main challenge towards definition of successful Data Spaces is the decision of what concrete standards and design principles are adopted, since they have to be accepted by all participants.

The following sections elaborate on the different components FIWARE brings materializing the different technical Building Blocks required for creation of Data Spaces. The i4Trust initiative is creating a curated framework based on these components that will be available for creation of Data Spaces and pioneers experiments during the period 2021-2023.

2.1

DATA INTEROPERABILITY BUILDING BLOCKS

Data providers joining Data Spaces must be able to publish data resources at well defined endpoints knowing that data consumers, a priori unknown to them, will know how to retrieve and

consume data through those endpoints. Data consumers, on the other hand, must know how data available through endpoints they discover can be consumed. This is a key principle which was observed in the design of the world wide web: content providers publish web pages on web servers (endpoints) knowing that web browsers will be able to connect to them and retrieve web pages whose content they can render and display to end users. It means that all participants in Data Spaces should 'speak the same language', which translates into adopting domain-agnostic common APIs and security schemas for data exchange (the way of constructing sentences) together with data models represented in data formats compatible with those APIs (the vocabulary used in constructed sentences).

The NGSI API is domain-agnostic. Actually, many different systems have been developed using NGSI in domains such as Smart Cities, Smart Manufacturing, Smart Energy, Smart Water, Smart AgriFood, Smart Ports, or Smart Health, to mention a few. This facilitates data sharing because each system participating in a Data Space will be publishing data that simply enriches a Digital Twin data representation of the world that the rest of systems connecting to the Data Space will know how to access. Systems participating into the Data Space don't know a priori what other systems may consume the data they publish (although they will be able to set up concrete terms and conditions for accessing/using data as we will explain in the next section).

Figure 6 illustrates how different systems participating in Data Spaces "powered by FIWARE" will exchange data. Context Broker servers are the endpoints through which systems connected to the Data Space publish Digital Twin data, very much like web servers publish html content on the world wide web. Those systems can in turn connect to Context Broker servers in order to obtain information they need. Note that Data Spaces powered by FIWARE enable near real-time (right-time) exchange of Digital Twin data which is fundamental in the design of innovative value chains demanding a very dynamic exchange of data among participants. Just think about scenarios like a city managing traffic lights in streets close to a given train station in order to facilitate that travelers arriving and taking a taxi can leave faster to their destinations. NGSI-LD brings very simple and therefore easy to use operations for creating, updating and consuming context / Digital Twin data but also more powerful operations like sophisticated queries, including geo-queries, or the subscription to get notified on changes of Digital Twin entities. On the other hand, Data Spaces "powered by FIWARE" can also support the exchange of large files using standard file transfer protocols, since this kind of file transfer may be required for certain scenarios like training of AI algorithms.

Note that systems participating in Data Spaces "powered by FIWARE" do not need to be themselves "powered by FIWARE". Systems which have not been architected using FIWARE can still use the NGSI API to share data they produce and consume data they need in the form of data associated with attributes of Digital Twin entities which represent that part of the world they deal with. This can be done directly by the systems or through NGSI system adapters which have been programmed to perform a conversion between NGSI and the API that the system natively supports for managing data.

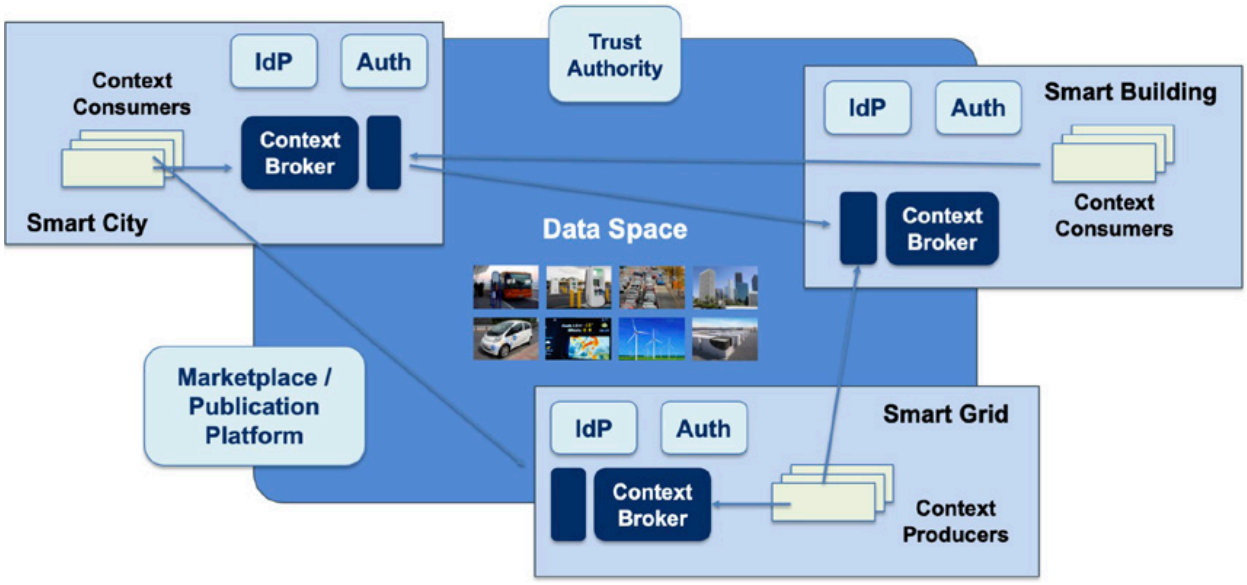

Figure 6 - Data Exchange in a Data Space "powered by FIWARE"

From a theoretical perspective, systems connected to a Data Space should be able to share data using the API they prefer. The specification (information model) of each API could be published as some kind of manifesto that certain components, integrated as part of the platform that systems should use to connect to the Data Space, can dynamically process in order to perform an automated adaptation from/to the APIs. However, such an approach faces important challenges. In the first place, such an approach has only been demonstrated in very simple scenarios and involving very simple APIs. Thus, the ability to exploit the kind of sophisticated features NGSI-LD would support for accessing Digital Twin data will be rather limited. On the other hand, creating a "common lingua" has proven to work for creating the kind of ecosystems we pursue: we shall ask ourselves if the world wide web had experienced the speed on adoption reached if HTTP and HTML hadn't been adopted as "common lingua" for web servers and browsers and each web server had the ability to choose a different protocol (as opposed to HTTP) or a different document format (as opposed to HTML). NGSI-LD has the advantages of being an open standard (defined by ETSI), has a strong Open Source community behind (FIWARE) and, quite relevant within Europe, it will pave the way for alignment with developments in the Connecting Europe Facility (CEF) program. Creation of system adapters that transform from specific APIs a given system may still require to use from/to NGSI-LD has proven to be not a complex task.

As mentioned before, the NGSI-LD API is domain-agnostic, therefore it is designed to work for any type of Digital Twin. Consequently, achieving full interoperability requires also the adoption of common data models to be represented in formats compatible with the APi. Here, the Smart Data Models initiative, already introduced

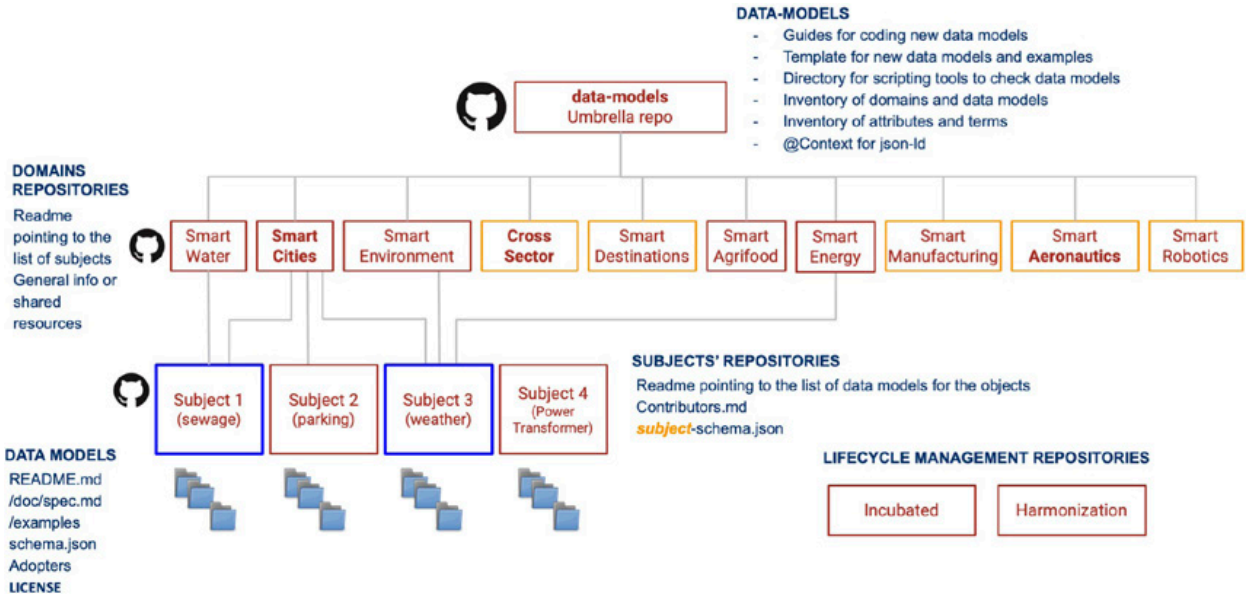

Figure 7 - Smart Data Models organization on GitHub

in a previous section, reaches great relevance. It brings a powerful resource for developers who can rely on the way data model specifications are mapped into concrete JSON and JSON-LD structures under the initiative, compatible with NGSIv2 and NGSI-LD respectively.

Figure 7 illustrates how resources are organized within the Smart Data Models initiative on GitHub. Data Models are grouped into "subjects" (weather, parking, aquaculture, etc) which in turn are referred to from repositories associated with the multiple application domains being considered (Smart Cities, Smart AgriFood, Smart Manufacturing, Smart Water, Smart Energy, etc). Note that there are subjects which are very specific to a given application domain (e.g., "street lighting" with regards to smart cities and communities) while others may be relevant to multiple domains (e.g., "weather" that is relevant to almost every domain or "sewage" that is relevant to the Smart Cities and Smart Water domains). Published data models are open to contributions and royalty-free.

An open governance model has been defined for the Smart Data Models initiative defining the lifecycle of data models comprising incubation of brand new data models as well as curation of data models via harmonization of different contributions. Processes and procedures for management of the different activities follow best practices from Open Source communities, guided by principles of transparency and meritocracy.

Completing the picture of Building Blocks for Data Interoperability, FIWARE brings components which provide the means for tracing and tracking in the process of data provision and data consumption/use. It provides the basis for a number of important functions, from identification of the provenance of data to audit-proof

logging of NGSI-LD transactions. For those Data Spaces with strong requirements on transparency and certification, FIWARE brings components (i.e., Canis Major) that ease recording of transaction logs into different Distributed Ledgers / Blockchains.

2.2

DATA SOVEREIGNTY AND TRUST BUILDING BLOCKS

Data Spaces must provide means for guaranteeing organizations joining Data Spaces that they can trust the other participants and that they will be able to exercise sovereignty on their data. That requires the definition of common Building Blocks, based on mature security standards that will be used by all participants in the Data Space.

A first fundamental building block to support within Data Spaces has to do with **Identity Management (IM)**. This building block allows identification, authentication, and authorization of organizations, individuals, machines and other actors participating in a Data Space. While this building block can be implemented on the basis of third party Open Source technologies like KeyCloak, just to mention one popular example, FIWARE brings the **Keyrock** component which supports OpenIdConnect, SAML 2.0 and OAuth2 standards. Quite relevant for Data Spaces deployed in Europe, Keyrock also resolves integration with eIDAS, a building block provided by the European Commission that enables the mutual recognition of national electronic identification schemes (eID) across borders, allowing European citizens to use their national eIDs when accessing online services from other European countries.

A second fundamental building block will be the one which facilitates trusted data exchange among participants, providing certainty that participants involved in the data exchange are who they claim to be, and that they comply with defined rules/agreements. Trust refers to the fact that data providers and data consumers can rely on the identity of the members of the data ecosystem and beyond that, between different security domains. Here, **IDS Connector technology**, as described in the IDS Reference Architecture Model (RAM) emerges as a solid basis and the FIWARE Community has incubated an Open Source implementation of this technology that has already been tested on how it can integrate with the rest of core FIWARE components.

Figure 8 illustrates how an AI service provider application and an AI service consumer application, each hosted in different organizations and participating in a common Data Space "powered by FIWARE", may interact with each other. It shows the role of components that are part of the incubated FIWARE IDS Connector implementation and their integration with other FIWARE components. The AI service provided can be an advanced traffic prediction service while the AI service consumer application might be a traffic management system running in a given city. The

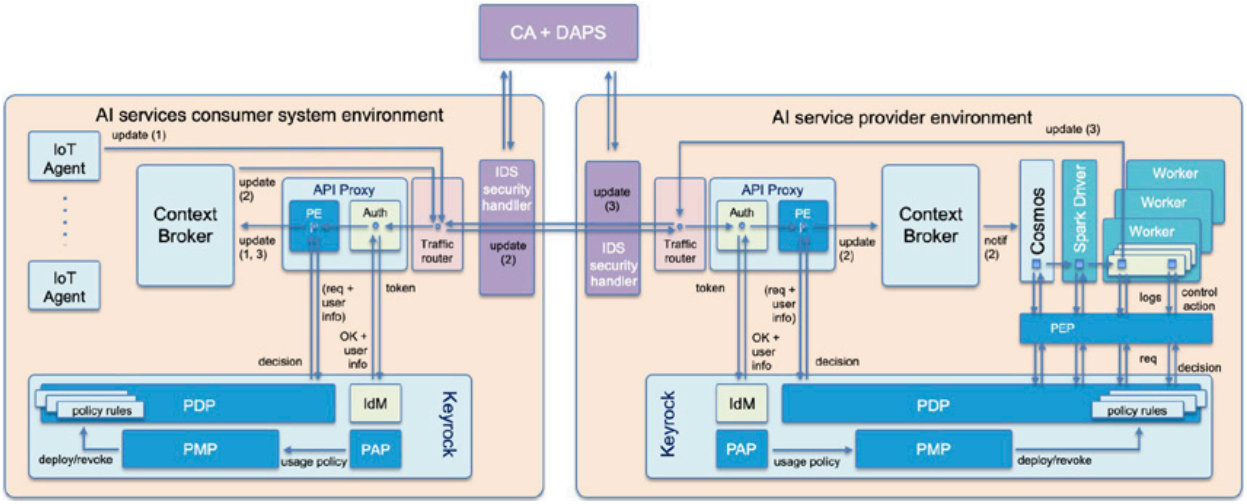

Figure 8 - Integration of FIWARE IAM and IDS Connector functionalities

AI service consumer wishes to incorporate near real-time traffic predictions based on current traffic levels measured through IoT sensors deployed in the streets. The Context Broker component running on the consumer application sends notifications based on updates on attribute "current traffic" of a given street. These updates are notified to the AI service where the value is processed in real-time applying AI algorithms (e.g. using Apache Spark). The predicted traffic in about 30 minutes is calculated (generating, for example, values "low", "medium" or "high") and the value of attribute "traffic in 30 minutes" of the given street is updated if changes occur.

In FIWARE, the implementation of IDS Connectors comes closely linked to the deployment of Kubernetes clusters. In this respect, the implementation of the concept of IDS connectors would not be limited to the implementation of the procedures used to exchange data among IDS connectors but also the monitoring and control of what is deployed in the associated clusters (e.g., guaranteeing that only certified components/applications are deployed) as well as the control of the flow of data within the cluster, ensuring enforcement of data access/usage control policies. For this last purpose, the usage of products supporting service mesh concepts like Istio are essential. In figure 8, such components are identified as "traffic routers".

In the IDS RAM, identity and access/usage control is currently performed at organization level. The **Certification Authority (CA)** and **Dynamic Attribute Provisioning Service (DAPS)** are in charge of verifying that organizations exchanging data are known and trustworthy. The IDS Connectors ensure that data exchanges are authorized at organization level. In figure 8, for example, the FIWARE IDS Connector running at the AI service consumer side ensures that when an IoT agent updates context / Digital Twin attributes on the local Context Broker (1) forwarding of updates (2) are sent out only to authorized organizations like the AI service provider in this

example. The FIWARE IDS Connector running at the service provider side will in turn verify that the organization from which updates are received is a trusted party and is authorized to make such updates to be then forwarded through the FIWARE Cosmos component to Spark where AI algorithms for traffic prediction purposes run. Similarly, when updates on the Context Broker deployed at the consumer side are invoked from the AI algorithms (3), the FIWARE IDS connector at the AI service provider side verifies that only requests authorized to be transmitted will be sent out. The FIWARE IDS connector at the AI consumer side verifies that the update request received comes from a trusted and authorized organization. All these verifications take place relying on the global CA+DAPS services of the Data Space.

On top of this authentication and data access/usage control procedures, performed at organization level via IDS Connectors, FIWARE security components enable an additional and finer grained authentication and access control at the end user level. Therefore, coming back to *Figure 8*, notifications received at the AI service provider side, same as update requests handled at the AI service consumer application side carry additional security tokens (namely, Json Web Tokens - JWT) linked to users on behalf of which those notifications/requests have been issued. The API proxy components, both at the consumer and on the provider side, validate those tokens and obtain the info associated with the corresponding users relying on standard Identity Provider services supported by the FIWARE Keyrock component. For managing access control, a standard XACML process is implemented. The API proxy plays the role of the Policy Enforcement Point (PEP) and the FIWARE **AuthZForce** component, also alternatively Keyrock, implements the Policy Decision Point (PDP) functionality. That means that the proxy forwards user info of the particular user together with information about the concrete notification/request to the PDP which then checks whether users with those credentials are entitled to perform the given operation. FIWARE brings alternative implementations of the API proxy, namely **Wilma**, **API Umbrella** and **CoatRack**, all of them compatible with Keyrock or any third party product implementing OpenID Connect, OAuth2 and XACML standards. Keyrock also implements Policy Administration Point (PAP) and Policy Management Point (PMP) standard XACML functions.

2.3

DATA VALUE CREATION BUILDING BLOCKS

Loose coupling of participants is a fundamental principle in Data Spaces. Data providers and consumers do not necessarily know about each other. Therefore, it becomes essential to incorporate Building Blocks enabling the management of **data resources as true assets** with a business value. Assets which can be published, discovered and, eventually, traded. This way boosting the

creation of multi-side markets where innovative services can be created.

FIWARE Business Application Ecosystem (BAE) components enable creation of **Marketplace services** which participants in Data Spaces can rely on for publishing their offerings around data assets they own. Different types of data assets can be defined via plugins but three kinds of standard data asset types are supported by default, namely static data files, right-time data resources provided via NGSI-LD at well defined endpoints, and data processing services, which typically have associated well defined endpoints for providing input data and publish results, both in right-time, using NGSI-LD. Marketplace services are accessible through a portal or via APIs. Users, either end users through the portal or applications via APIs, of the Marketplace service can perform the following main actions:

- define new data asset types via plug-ins;
- register offerings around defined data asset types which typically means providing description of the asset, data models behind, endpoints (URLs) through with the data asset will be accessible and, very important, terms and conditions defined around the data asset, including SLAs, legal clauses, access control policies users of the asset have to comply with and associated pricing schema, which may be based on:
  - Free access (open data): user pay no money;
  - One-time payments: users pay only once;
  - Recurring payments: users pay periodically (monthly, yearly, etc) for getting access to some data. In addition, users will be able to cancel the subscription, but they won't be able to access data anymore;
  - Usage payment: users pay per use. Their payments are based on the amount of information consumed.
- ability to navigate and search/discover existing offerings based on selected criteria.

Following is the list of backend components and APIs associated to the FIWARE BAE Marketplace:

- Backend implementing standard TM Forum APIs supporting configuration of the marketplace:
  - Catalog Management API
  - Product Ordering Management API
  - Product Inventory Management API
  - Party Management API
  - Customer Management API
  - Billing Management API
  - Usage Management API
- Rating, Charging, and Billing backend;
- Revenue Settlement and Sharing System;
- Authentication, API Orchestrator, and Web portal.

Figure 9 shows the FIWARE Data Marketplace portal deployed and configured for the i4Trust project.

FIWARE also comprise components for publication of data resources linked to data assets around which offerings are managed through the FIWARE Data Marketplace. For this purpose, we have the Idra publication platform and an extended version of the CKAN open data platform, which is an open data publication platform widely adopted in the market. These extensions support enhanced data management capabilities and integration with FIWARE technologies including NGSI. In particular, data

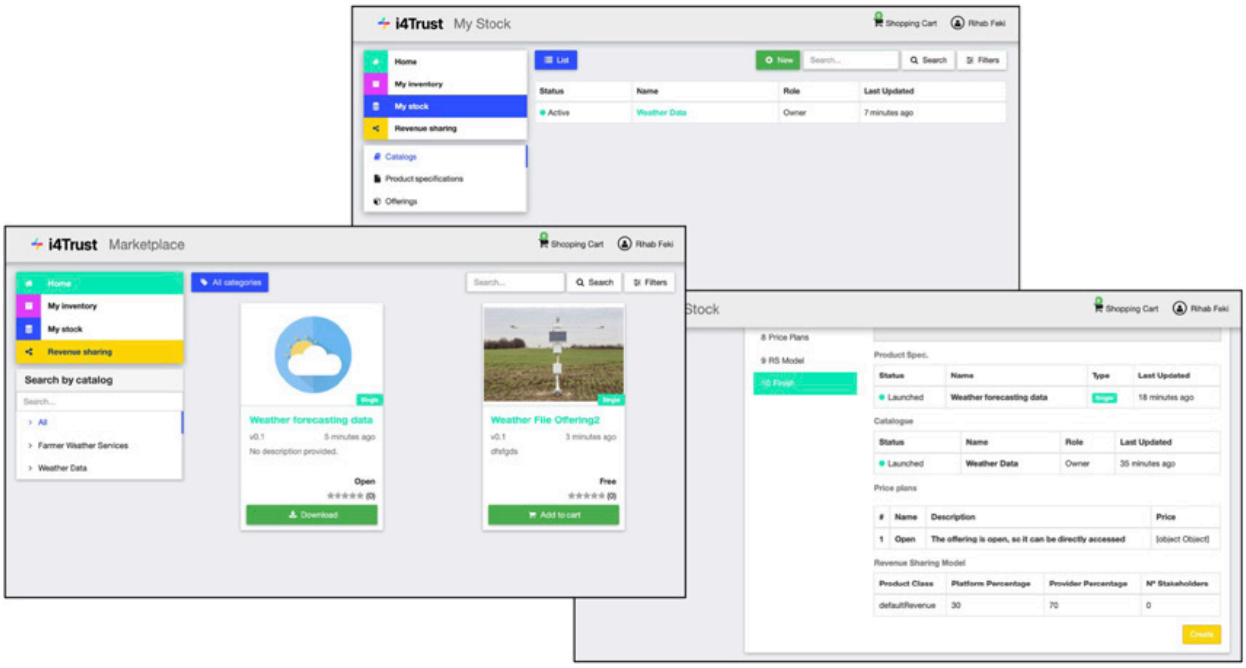

Figure 9 - FIWARE Data Marketplace portal

publication and discovery features provided by CKAN have been enhanced with the following features:

- Right-time (near real time) time data publication - thanks to this extension, CKAN is not limited to list data resources linked to static files as part of its catalogue but also data resources linked to NGSI-LD requests served by Context Broker components deployed in a Data Space. This brings the ability to discover data resources relying on DCAT capabilities publication platforms support;
- Identity Management, Authentication and Access Control functions based on Keyrock components - therefore supporting OpenId Connect, OAuth2 and XACML standards adopted at overall Data Space level;
- Publication of priced data resources - thanks to this extension, it is possible to mark data resources listed as part of the catalogue as linked to offerings visible in the Data Marketplace. Users can therefore click on those data resources and navigate to the Marketplace to proceed with the acquisition of access rights;
- Enhanced Data Visualization - thanks to this extension WireCloud (Configurable Dashboards component) to create adaptive and incremental data visualization. This way, farmers (sellers) can decide the way they want their data to be shown.

3 Towards development of European Data Spaces

3.1

ALIGNMENT WITH CONNECTING EUROPE FACILITY (CEF) PROGRAM

The Connecting Europe Facility (CEF) Digital programme, created as part of the CEF Telecom program, is aimed at pursuing real improvements in daily life for citizens, businesses (including SMEs) and administrations through the deployment of solid trans-European Digital Service Infrastructures (DSIs) based on mature technical and organisational Building Blocks which can be used by any public or private organisation. CEF Digital focuses on providing operational services which are ready to be deployed out-of-the-box and which will be both sustainable and maintained over the long term. Both legal texts that articulate the CEF Telecom programme - the CEF Regulation and CEF Telecom Guidelines - repeatedly recognise the contribution of the programme to the achievement of the Digital Single Market. As an ongoing task, it is intended that the use and processing of data at all levels is improved thanks to deployment of CEF DSIs (based on CEF Digital Building Blocks) and that this will enable the creation of solutions that are replicable across member states and interoperable cross-frontier. This initiative will effectively support the completion and sustainability of an internal Digital Single Market that will enhance the competitiveness of the European Economy and will push European society forward.

The European Commission (EC) has formally adopted FIWARE Context Broker technology as CEF Building Block within their Digital CEF (Connecting Europe Facility) Building Blocks program. On the other hand, the FIWARE Keyrock component has been successfully integrated with the eID CEF Building Block, which brings the means for electronically identifying users (organizations or individuals) from all across Europe. Last but not least, FIWARE incorporates components that help to define, by configuration without coding, what concrete logs on Context Broker transactions (e.g., updates, queries, notifications, etc) to store into blockchains or distributed ledgers. Such components will support integration with the European Blockchain Service Infrastructure (EBSI), incorporated as part of the CEF Building Blocks in 2019.

This means that the proposed vision for Data Spaces will bring as benefit a strong alignment with the CEF Digital vision and strategy when it comes to the European Union.

3.2

MATERIALIZING EUROPEAN DATA SPACES

In February 2020, the European Commission announced the European Strategy for Data, aiming at creating a single market for data to be shared and exchanged across sectors efficiently and securely within the EU. Behind this endeavour stands the Commission's goal to get ahead with the European data economy in a way that fits European values of self-determination, privacy, and fair competition. For this to achieve, the rules of accessing and using data must be fair, clear and practicable. This is especially important as the European data economy continues to grow rapidly – from 301 billion euros (2,4 % of GDP) in 2018 to an estimated 829 billion euros (5,8 % of GDP) by 2025.

The centrepiece of the European Data Strategy is the concept of "Data Spaces", for which the Commission defined nine initial domains, all driven by sector-specific requirements. Actually, the Commission promotes the development of European Data Spaces for strategic economic sectors and public-interest domains, starting with the following nine: industrial (manufacturing), green deal, mobility, health, financial, energy, agriculture, public administration, and skills.

While Data Spaces stimulate higher availability of data pools, technical tools, and infrastructure addressing domain-specific challenges and legislations, the EU Strategy for Data acknowledges that these Data Spaces should be interconnected and that this challenge requires specific attention. But Europe doesn't need to start from scratch – data sharing and exchange within specific domains and sectors is already happening in existing initiatives. However, each of these initiatives follows its own approach, and therefore they are not always interoperable.

Just like other technology infrastructures (e.g. the internet), Data Spaces basically are sector-agnostic, with many requirements and functions being similar or even identical across different sectors and Data Spaces. Therefore, creating the basis for Data Spaces primarily is not so much a technological challenge, as there are plenty of technical solutions and standards available. The main challenge towards interoperable Data Spaces is to agree on open standard-based Building Blocks and design principles that are accepted by all participants. While making it work in pilot applications, proof of concepts, and living labs is relatively easy, the real challenge lies in the convergence of interoperability to new norms and allowing for mass adoption and scalability. Alignment with the CEF (Connecting Europe Facility) Digital program would be highly desirable, since this program is precisely devoted to bring the Building Blocks for creating Digital Service Infrastructures in Europe.

On the other hand, GAIA-X is aimed at creating a federated form of data infrastructure in Europe which strengthens the ability to both access and share data securely and confidently. Since its creation, the initiative has raised a lot of aware-

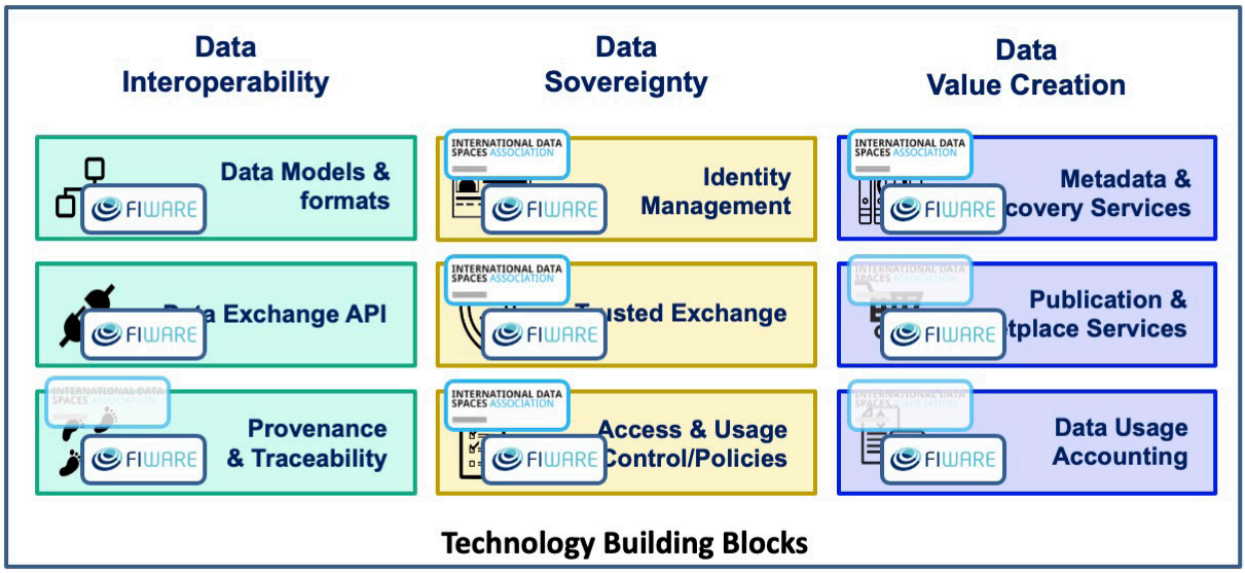

Figure 10 - IDSA and FIWARE positioning to create Data Spaces*

ness because of the opportunity it brings to join forces and, leveraging existing assets, accelerate delivery of solutions to the market.

As explained earlier, IDSA is providing several architecture elements which are necessary to create Data Spaces with trust and data sovereignty being some of the most important ones. These are necessary but not sufficient to create real living Data Spaces which are accepted and adapted on the market. Standard API's, standard data models as well as marketplace functionalities are also necessary and this is where the core strength of the FIWARE ecosystems lies. The complementarities of both initiatives regarding technology Building Blocks is illustrated in *Figure 10*. As the FIWARE Context Broker technology was adopted in 2019 as one of the Building Blocks of the Connecting Europe Facility (CEF) program and has proven to integrate smoothly with other relevant CEF Building Blocks for the creation of Data Spaces (e.g., eIDAS, EBSI) this would bring an opportunity of alignment with this important program. As a result, the close partnership between FIWARE and IDSA is guaranteeing the aligned contribution of both ecosystems to the success of Data Spaces based on GAIA-X.

* Labeling of a building block with the IDSA logo means that IDSA has produced or is working on specifications. In some cases, specifications are in an early stage. Labeling with the FIWARE logo means that the FIWARE Community has produced Open Source implementations of components and, eventually, is driving standardization activities. When both logos are present, an opportunity of collaboration has been identified linked to evolution of the specifications following an Open Source implementation approach.

Conclusions

FIWARE brings the essential Building Blocks helping to create Data Spaces enabling access and share of data in an effective and trustworthy manner. The Open Source nature of FIWARE technologies foster creation of these Data Spaces as federated infrastructures where multiple providers can be involved and not just a few players. FIWARE is bringing today mature technologies, compatible with IDS and CEF Building Blocks, that may accelerate the delivery of Data Spaces to the market and the materialization of initiatives like GAIA-X in Europe. In addition the FIWARE ecosystem has proven the ability to create standards which are accepted and adapted in the meantime on a global scale. Actually, solution providers and end users in the Americas, in Africa, in India or in Asia are building their smart solutions based on FIWARE standards and technologies. Relying on FIWARE, there is the same potential to create solutions and to define standards for Data Spaces which can be adopted also on a global scale.

FIWARE
Open APIs for Open Minds

www.fiware.org

FIWARE FOR DATA SPACES

POSITION PAPER

VERSION 1.0 / JUNE 2021