Alliance for Internet of Things Innovation

Guidance for the Integration of IoT and Edge Computing in Data Spaces

Release 1.0

AIOTI WG Standardisation
Task Force High Level Architecture

23 September 2022

Executive Summary

This document provides an analysis on the integration of IoT and edge computing in data spaces.

It explains the context, providing a definition of data spaces, enumerating challenges of data spaces, as well as the positioning of data spaces in the AIOTI high-level architecture (HLA) [1].

It provides an architecture analysis of data spaces, covering:

- data space systems of interest from three perspectives: computing continuum, federation of systems, and data collecting / trading;
- stakeholders of data space systems;
- concerns and properties, general to data spaces, specific to cyber physical systems, to the integration of edge computing and processing, and to trustworthiness;
- building blocks to address concerns related to data governance, cyber physical systems and digital twins, trustworthiness support, interoperability support, infrastructure reconfiguration support, and data business marketplaces.

It describes the relation to existing solutions:

- a construction approach relying on reference architecture standards and patterns:
- the use of reference architectures proposal from IDSA, oneM2M, ETSI MEC;
- the work carried out by a number of large-scale projects: PLATOON, INTERCONNECT, SmartBear, ASSIST-IoT.

It provides recommendations for data space standards.

[1] https://aioti.eu/wp-content/uploads/2020/12/AIOTI_HLA_R5_201221_Published.pdf

Table of Content

Table of Figures

List of Tables

Context

Data Spaces

While the term **data space** was coined more that 10 years ago[2], it was not until recent years that a number of position papers such as BDVA[3,4], OpenDei[5], and initiatives, such as IDSA[6], or GAIA-X[7,8] or FIWARE[9] have started to propose a common understanding.

OpenDei provides a comprehensive definition:

> *From a technical perspective, a **data space** can be seen as a data integration concept which does not require common database schemas and physical data integration, but is rather based on distributed data stores and integration on an "as needed" basis on a semantic level. Abstracted from this technical definition, a data space can be defined as a federated data ecosystem within a certain application domain and based on shared policies and rules*

FIWARE provides a definition which is aligned:

> *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.*

In this position paper, we will assume that a data space is a **trustworthy decentralized environment for data sharing**.

Decentralisation is a particularly important characteristic as showed in Figure 1. It provides an example of data spaces with five organisations engaged in carrying out operations on data. The figure highlights

- two layers: the processing layer, and the data layer; and
- three concepts:
    - data exchanges: a relationship that involves organisations,
    - data interoperability: a capability between processing systems, and
    - data operations: activities carried out by processing systems.

---

[2] https://en.wikipedia.org/wiki/Dataspaces
[3] Towards a European-Governed Data Sharing Space. Enabling data exchange and unlocking AI potential. April 2019 https://bdva.eu/sites/default/files/BDVA%20DataSharingSpace%20PositionPaper_April2019_V1.pdf
[4] Towards a European-Governed Data Sharing Space. Enabling data exchange and unlocking AI potential. November 2020 https://www.bdva.eu/sites/default/files/BDVA%20DataSharingSpaces%20PositionPaper%20V2_2020_Final.pdf
[5] https://design-principles-for-data-spaces.org/
[6] https://www.internationaldataspaces.org/wp-content/uploads/2019/03/IDS-Reference-Architecture-Model-3.0.pdf
[7] https://www.data-infrastructure.eu/GAIAX/Redaktion/EN/Publications/gaia-x-technical-architecture.pdf?__blob=publicationFile&v=5. Release-June 2020
[8] https://www.data-infrastructure.eu/GAIAX/Redaktion/EN/Publications/gaia-x-technical-architecture.pdf
[9] https://www.fiware.org/marketing-material/fiware-for-data-spaces -(release June 2021)

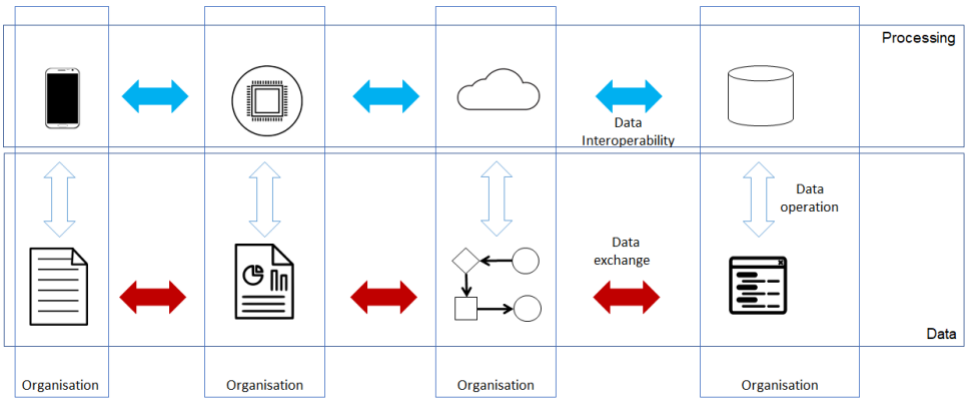


Figure 1 – Decentralised data space example

The two layers (processing layer, data layer) and the three concepts (data exchange, data interoperability and data exchange) can be used to illustrate data spaces in various configurations. Figure 2 provides an example illustrating AI capability in a digital twin:

- the processing layer focuses on knowledge handling and reasoning, while the data layer focuses on and knowledge representation and storage;
- data exchange takes place between the handling and reasoning capabilities of the virtual entity and the same capabilities of the physical entity;
- data interoperability is enabled by knowledge representations agreed between the virtual and the physical entity. and
- data operations are carried out handling and reasoning capabilities of the virtual entity and the same capabilities of the physical entity.

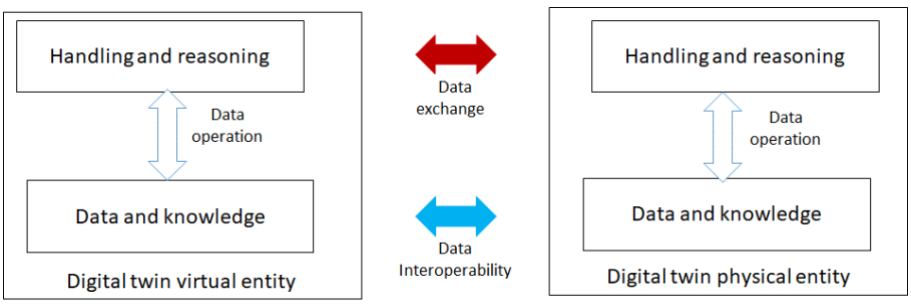


Figure 2 – AI capability in a digital twin example

Data Spaces Principles and Challenges

The advent of data spaces will depend on whether we are able to apply a number of principles in order to solve a set of challenges. They are described in the table below.

Table 1 – Data space principles and associated challenges

Principles	Challenges	Description
Data spaces are ecosystems of systems	Structuring and operating an ecosystem of ecosystems	Technology ecosystems (e.g., 5G, Clouds, IoT and Edge, AI) must be combined with vertical domain specific ecosystems (e.g., smart manufacturing, health, energy, agriculture). The structure of a resulting ecosystem of ecosystems has to be created, concerning services and infrastructure, stakeholders and orchestration of ecosystems.
Data usage require provisioning from connecting devices	Creation of value associated with usage control	The trend today for IoT connectivity creates the potential for an economy based on extensive data usage. The vision of free flow of data must also be associated with data sovereignty, usage policies and trust. Access control to data is not sufficient, it must be replaced by usage control of data at the business level as well as at the consumer level. Usage control at the business level can involve IPR considerations, or regulatory considerations. Usage control at the consumer level can involve privacy considerations.
Data spaces support data lifecycle	Characterizing and managing data lifecycle	Data spaces must support the entire data lifecycle, which can include the following stages: data concept, data requirements, data planning, data acquisition, data preparation, building model, system development, system operation, data decommissioning, system decommissioning.
Data interoperability enabled by a common language	Common language for semantic interoperability	A common language for data Interoperability and discovery is required. It requires the exchange of metadata based on ontologies and semantic information. It is used for knowledge discovery in the data space.
Data usage enabled by common data models	Common data models for behavioral interoperability	Common data models are required to ensure that data operations are consistent even though they can be carried out by different organisations. They should be domain agnostic and use representation formats that allow for exchange through APIs.
Data curation	Organisation, description, cleaning enhancing and preserving for public use	Data curation is important to maintain the value of data. A suitable data curation network and practice should available so that data can be organised, described, cleaned, enhanced and preserved for public use.
Trust in data sharing	Trustworthiness and risk management	Trustworthiness is an important concern in data sharing. It includes quality attributes such as privacy, transparency, accessibility, fairness, accountability, security. It also includes capabilities such as consent management or control of personal data.

Principles	Challenges	Description
		It is contributed by appropriate risk management, federated security management, federated privacy management and federated assurance management.
Governance for ethical usage of data	Governance and ethics	A suitable governance model should be applied spelling out clearly rights and responsibilities (e.g., what actions can be taken, by whom, with what data), with the capability to monitor compliance to decided policies. Such policies should include ethical considerations. The monitoring of the data space should be possible and key performance indicators can be available.
Decentralisation	Decentralisation	A decentralised architecture is needed. It has to be agreed upon by all relevant stakeholders of the ecosystem. This implies proper identity management and the use of common distributed agreement schemes (e.g., distributed ledger technologies).
Integrated data management	Data fabric	An integrated data management platform that enables the full breadth of integrated data management capabilities including discovery, federated governance, curation, and orchestration.
Extensible data spaces	Scaling-up data spaces	To enable the scale-up of data spaces, a virtual continuum must be created in the space and time dimensions. The space dimension focuses on enabling data exchange across the range of processors (IoT, Edge, Cloud), or across the range of systems (e.g., a smart solution in energy, a system integrated into the energy system of systems, or the federation or ecosystems, smart building, smart grid, smart mobility). The time dimension focuses on enabling the evolution of data spaces. It can cover entities of interest (e.g., data related to a vehicle, data related to a fleet of vehicles, or data related to a smart city transportation system).
User-centricity	Business roles and interactions	The creation of a data economy requires understanding of business roles and business interactions. Roles can include processing entities (e.g., data scientists), data providers (e.g., operators of sensing systems), data owners (e.g., consumers), and marketplace operators. The ecosystem must be user-friendly, support ownership enforcement, and provide room for consumer, business and public functionalities.

Data Spaces in the AIOTI High-Level Architecture

AIOTI has defined an architecture specification called HLA (for high-level architecture)[10], to support the work of IoT Large Scale Pilots (LSP) and to help them promote architecture building blocks. The HLA is based on three layers as defined in Table 2. Note that the term layer is used here in the software architecture sense. Each layer simply represents a grouping of modules that offer a cohesive set of services; no mappings to other layered models or interpretation of the term should be inferred.

Table 2 – AIOTI HLA layers

Application layer	Group application entities as well as communications and interface methods used in process-to-process communications
Intermediate layer or Data layer	Groups IoT specific functions, such as data storage and sharing, and exposes those to the application layer via interfaces commonly referred to as Application Programming Interfaces (APIs). The Data layer makes use of the Network layer's services Note that in the HLA specification, the Data layer is called IoT layer. We have chosen not to use the term IoT layer in order to avoid confusion with IoT devices.
Network layer	Services of the Network layer can be grouped into data plane services, providing short range as well as long range connectivity and data forwarding between entities, and control plane services such as location, device triggering, QoS or determinism. The following terms are used: - Plane: abstract conception of where certain processes take place[11] - Data plane: part of a network that controls how data packets are forwarded - Control plane: responsible for the forwarding of packets.

[10] https://aioti.eu/wp-content/uploads/2020/12/AIOTI_HLA_R5_201221_Published.pdf

[11] https://www.cloudflare.com/learning/network-layer/what-is-the-control-plane/

Figure 3 shows the data space example described in Figure 1 using the HLA representation. The difference is the addition of the network layer which puts emphasis on interoperability properties.

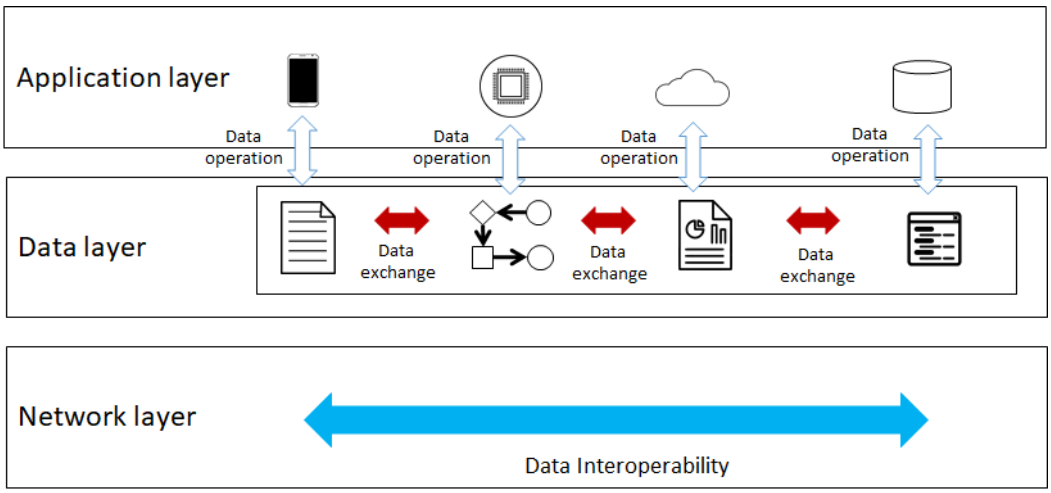


Figure 3 – Data space example using the HLA representation

Figure 4 shows the digital twin example described in Figure 2 using the HLA representation.

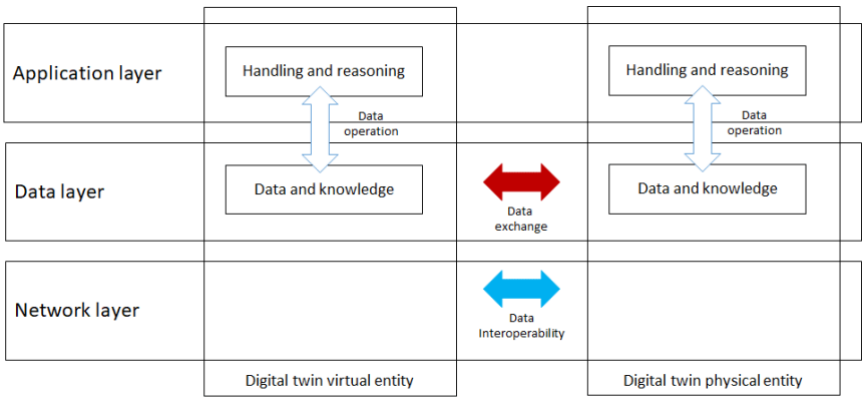


Figure 4 – HLA representation of digital twin example

1 IoT and Edge Computing Data Space Reference Architecture

1.1 Systems-of-interest

This section provides three important architecture perspectives: computing continuum, federation, and marketplace.

1.1.1 Computing Continuum Perspective

A computing continuum perspective integrating IoT and edge computing is needed. Figure 5 shows a data spaces where this continuum is visualised from left to right:

- IoT devices carry out some data operations and exchange data,
- Edge systems carry out further data operations and exchange further data,
- Cloud systems carry out further data operations and exchange further data

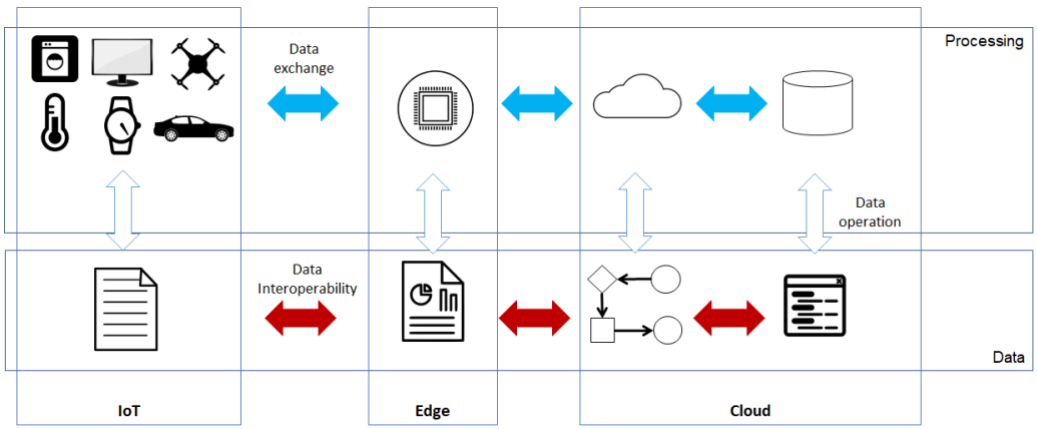


Figure 5 – Computing continuum perspective of data spaces

Figure 6 shows the same computing continuum perspective using the HLA representation

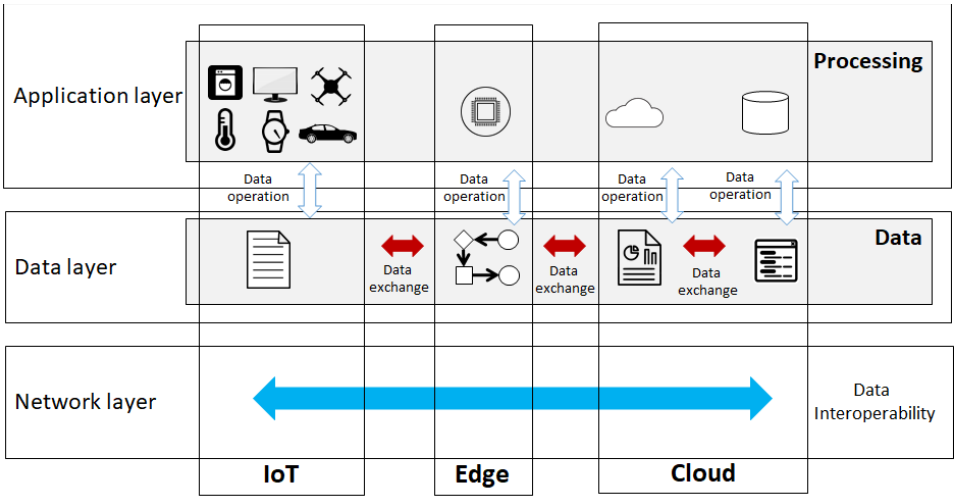


Figure 6 – Computing continuum perspective of data spaces based on HLA

1.1.2 Federated Systems Perspective

A federated system perspective can also be needed. Figure 7 shows this: while data exchange can take place within a data space ecosystem, two separate ecosystems can also exchange data. Federation is suitable in particular to achieve cross domain exchange e.g. between the energy and the transport domain as shown in figure 8.

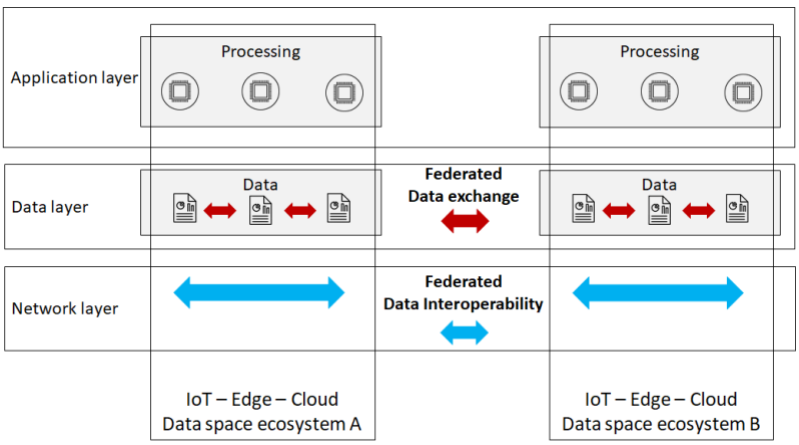


Figure 7 – Federated systems perspective of data spaces

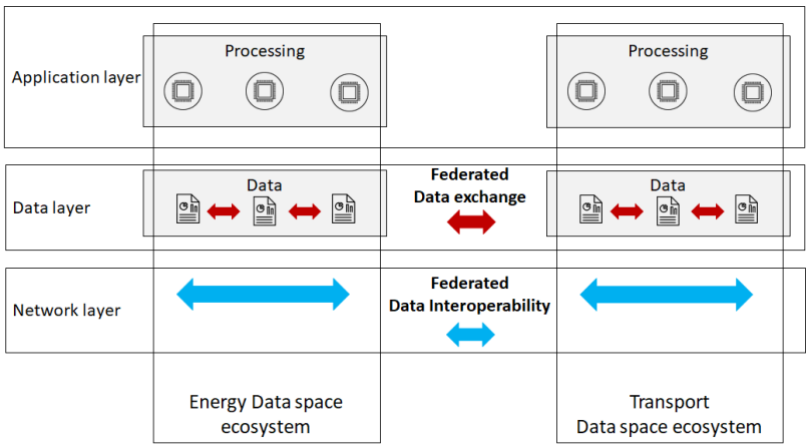


Figure 8 – Domain perspective of data spaces

1.1.3 Data Collecting and Trading Perspective

A data marketplace perspective can also be needed. Figure 9 shows a data collecting system, a data trading system, consisting of a market place, data providers and data consumers.

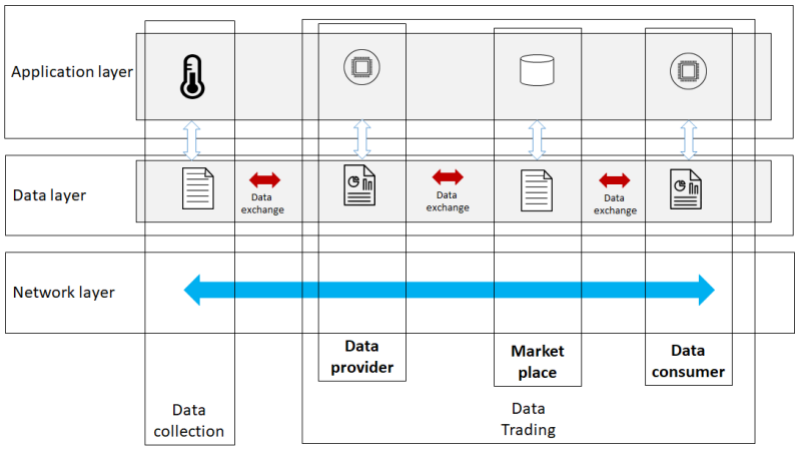


Figure 9 – Data collecting system and data marketplace

1.2 Stakeholders

The OpenDei position paper[12] lists the following stakeholders for data spaces:

- data consumers: they access data spaces to use data;
- data providers: they collect and manage data and make it available in data spaces;
- data producers: they create data;
- data owners: they have rights to grant or revoke terms and conditions for access and use of data;
- data application providers: they provide applications that transform, process or visualise data;
- data platform providers: they provide capabilities that allow for operation of data platforms;
- data marketplace providers: they provide capabilities that allow for operation of data marketplaces;
- identity providers: they provide capabilities for identifying parties.

The FIWARE Foundation's position paper[13] identifies the following stakeholders.

- Data Space Governance: they are responsible for managing the Data Space and ensuring that Data consumers and Data Owners are served as per the requirements
- Data Consumers: they are data consumers of the data in the Data Spaces. Typically the data aggregators and developers use the data to develop various solutions.
- Trust Providers: they provide centralised trust to Data Spaces and Data Space's data consumers.
- Data Owners: they own the data in the Data Space, and they should have sovereign authority over the data and its accessibility options

ISO/IEC 30141 (IoT Reference Architecture)[14] identifies the following stakeholders

- IoT service providers: they manage and operate IoT services. They can also provide network connectivity;
- IoT service developers: they implement, test and integrate IoT services with the underlying IoT platform
- IoT users: they use IoT services. There are both human users and digital users.

[12] https://design-principles-for-data-spaces.org/
[13] https://www.fiware.org/marketing-material/fiware-for-data-spaces
[14] https://www.iso.org/standard/65695.html

ISO/IEC 30164 (Edge computing)[15] identifies the following stakeholders

- Developers: they develop applications and services for the edge computing system;
- Service providers: the undertake business activities using the edge computing system;
- Equipment manufacturers: They produce devices used in edge computing (e.g., edge devices, IoT gateways, data centres);
- Consumers: they purchase edge computing devices and related devices for their own personal use.

Data space ecosystems which integrate IoT and Edge computing systems involve all these stakeholders as shown in the table below.

Table 3 – IoT and edge computing stakeholders

Data space stakeholders	IoT stakeholders involved	Edge computing stakeholders
Data consumers	IoT users	
Data providers	IoT service developers	Developers
Data producers	IoT service developers	Developers
Data owners		
Data application providers	IoT service providers	Service providers
Data platform providers	IoT service developers	Service providers and Equipment manufacturers
Identity providers	IoT service developers	Service providers and Equipment manufacturers
Data marketplace providers		

[15] https://www.iso.org/standard/53284.html

1.3 Concerns and Properties

1.3.1 Global concerns for data spaces

The OpenDei position paper[16] lists the following concerns concerning data spaces

- efficiency of data exchange, achieved through a suitable framework involving APIs, security schemes, and data models;
- agreement support and enforcement in data marketplaces;
- trustworthiness of the environment based on common ethical values where data consumers and data providers can engage into businesses; and
- policies and regulations support, through appropriate organisational and technical capabilities.

These concerns match the data space challenges presented in section 2.2 as follows:

Table 4 – Mapping between data space concerns and challenges

Data space concerns	Data spaces challenges (section 2.2)
Efficiency of data exchange	Ecosystem of ecosystems Scaling-up data spaces Data lifecycle Common language for semantic interoperability Common data models for behavioural interoperability Data curation Decentralisation
Agreement support and enforcement	Business roles and interactions
Trustworthiness based on common ethical values	Creation of value associated with usage control Trustworthiness and risk management
Policies and regulation support	Governance and ethics

[16] https://design-principles-for-data-spaces.org/

1.3.2 Global concerns for cyber physical systems

The Framework for cyber-physical systems (NIST special publication 1500-201, June 2017)[17] describes categories of concerns as shown in the table below.

Table 5 – Cyber physical systems concerns

Category of concern	Description	List of concerns
Functional	Concerns about function including sensing, actuation, control, communications, physicality, etc.	Actuation, communication, controllability, functionality, manageability, measurability, monitorability, performance, physical, physical context, sensing, states, uncertainty
Business	Concerns about enterprise, time to market, environment, regulation, cost, etc.	Enterprise, cost, environment, policy, quality, regulatory, time-to-market, utility
Human	Concerns about human interaction with and as part of a CPS	Human factors, usability
Trustworthiness	Concerns about trustworthiness of CPS including security, privacy, safety, reliability, and resilience.	Privacy, reliability, resilience, safety, security
Timing	Concerns about time and frequency in CPS, including the generation and transport of time and frequency signals, timestamping, managing latency, timing composability, etc.	Logical time, synchronisation, time awareness, time-interval and latency
Sovereignty	Concerns from the data owners about losing the control of their own data	Losing control of own data, Notion that once the data of data owner is on Dataspace then it is out of the data owner's control
Data	Concerns about data interoperability including fusion, metadata, type, identity, etc.	Data semantics, identity, operations on data, relationships between data, data velocity, data volume
Boundaries	Concerns related to demarcations of topological, functional, organisational, or other forms of interactions.	Behavioural, networkability
Composition	Concerns related to the ability to compute selected properties of a component assembly from the properties of its components. Compositionality requires components that are composable: they do not change their properties in an assembly. Timing composability is particularly difficult.	Responsibility, adaptability, complexity, constructivity, discoverability
Lifecycle	Concerns about the lifecycle of CPS including its components.	Deployability, disposability, engineerability, maintainability, operability, procureability, producibility

[17] https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1500-201.pdf

1.3.3 Integration concerns for edge computing and processing

The AIOTI paper on enabling technologies and challenges[18] lists emerging topics which can be considered as integration concerns as shown in the table below.

Table 6 – Integration concerns for edge computing

Category of integration concern	Comment
Digital twin integration	Digital twins provide an unifying view that has to be supported in Edge computing and processing
Deep Edge, Terminal and IoT Device Integration in B5G communication infrastructure	This includes the need for - a delivery model and APIs; - secure management of billions of devices; - privacy and data management; and - policy descriptions.
Edge, Mobile Edge Computing and Processing	This includes the need for - decentralised processing and storage to allow for federation and avoid centralised computation; - management of constraints on bandwidth, network congestion, latency, storage; - mobility capability (location and roaming support); and - edge-aware design approaches which avoid vendor lock-in, integrate operation constraints and get better availability
Network and Server security for edge and IoT	This includes the need for - addressing system-wide security challenges (lifecycle orientation, control and management of security, multi-tenancy, network virtualization and slices, edge and IoT); - addressing operational security capability (security quantification, green security, security as a service, security orchestration, disruptive security strategies, DLTs, AI and human-centric privacy)
Plug and Play Integrated Satellite and Terrestrial Networks	This includes the need for - diversification of the spectrum usage across multiple technologies; - edge networks to reduce the impact of the backhaul in the end-to-end system; - adapted data path protocols to massive communication environments; - application protocols adaptation through the virtualization environment; and - addressing the M2M communication needs in an efficient manner Participation within the main standardisation organisations: 3GPP, ETSI NFV, ETSI MEC, IETF, ONF
Autonomous and Hyper-connected On-demand Urban Transportation	This is based on the so-called Collaborative, Connected and Automate Mobility (CCAM) paradigm. It includes the need for - complex data management and analysis systems and infrastructures; and - tracking of the history of vehicles for maintenance of compliance to regulation purposes possibly through distributed ledger technologies.
Large scale deployment of IoT systems	It includes the need for - support of multiple connectivity solutions; and - energy awareness to take advantage of edge-enable energy resources

[18] https://aioti.eu/wp-content/uploads/2021/10/AIOTI-Beyond-5G-R1-Report-Published.pdf as well as https://bscw.5g-ppp.eu/pub/bscw.cgi/d367342/Networld2020%20SRIA%202020%20Final%20Version%202.2%20.pdf

1.3.4 Impact of the Trustworthiness Concern: Full stack integrity

Trustworthiness is defined as "the ability to meet stakeholders expectations in a verifiable way"[19]. It is a combination of attributes such as reliability, resilience, security, privacy, safety, availability, transparency, usability. The way these attributes are combined depend on the domain and the requirements of the underlying applications.

The following trustworthiness concerns can be raised in systems of systems:

- ensuring trustworthiness from a system of systems viewpoint, i.e., in a context where there are multiple systems; and
- ensuring trustworthiness from an interoperability viewpoint, i.e., in a context where one system has to interact and interoperate with another system;

The following trustworthiness concern can be raised in data space systems of systems:

- Ensuring trustworthiness of usage along the stack of systems involved in data exchange, as shown in Figure 10. We call this concern the full stack data usage integrity concern.

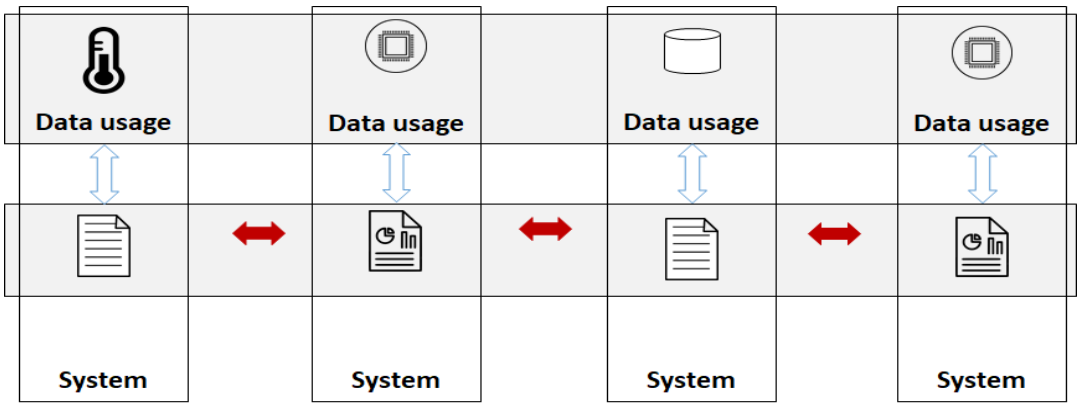


Figure 10 – Full stack data usage integrity

1.4 Building Blocks to Address Concerns

This section lists important building blocks that help address the listed concerns.

1.4.1 Data Governance Building blocks

A number of terms are used by ISO 8000-2 (Data quality - Part 2: Vocabulary) and ISO/IEC 38500 (Governance of IT for the organisation) as shown in the table below.

Table 7 – Data governance terms

[19] ISO/IEC TR 24028:2020 Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence

Term	Definition	Reference
Data governance	Development and enforcement of policies related to the management of data	ISO 8000-2 (Data quality - Part 2: Vocabulary)
Information technology (IT)	Resources used to acquire, process, store, and disseminate information	ISO/IEC 38500 (Governance of IT for the organisation)
Organisational governance	System by which organisations are directed and controlled	
Policy	Intentions and direction of an organisation as formally expressed by its governing body or executive managers acting with appropriate authority	

ISO/IEC 38500 further defines six principles of information technology governance which also apply to data governance:

- Responsibility: evaluate the options for assigning responsibilities on data, direct the responsibilities, and monitor them.
- Strategy: evaluate the need for business processes on data, direct strategies, and monitor them
- Acquisition: evaluate the need for data capabilities, direct their acquisition, and monitor their deployment
- Performance: evaluate the performance of data capabilities, direct their use, and monitor their deployment,
- Conformance: evaluate the extent to which data capabilities satisfy obligations, direct their enforcement and monitor compliance.
- Human behaviour. Evaluate human behaviour and take them under consideration, direct capabilities to be consistent with human behaviour, and monitor consistency of data activities.

Data governance therefore need support for responsibility, strategy, acquisition, performance, conformance and human behaviour evaluation, direction and monitoring

Data governance building blocks should consider the following dimensions:

- the policy dimension (e.g., policy on data access, policy on data usage. These policies enable data sovereignty);
- the enforcement across the continuum infrastructure, including at IoT, edge and cloud level, and
- the enforcement across the ecosystem, i.e. organisational governance.

1.4.2 Cyber Physical System and Digital Twins support building blocks

As shown in figure 5 and figure 6, a global infrastructure for data space includes IoT and edge capabilities. In these configurations, data spaces include cyber physical systems. In an overall federated data space environment, multiple cyber physical systems might coexist, e.g., energy systems, transport systems, health systems, manufacturing systems, agriculture systems, maritime systems and so forth.

Consequently data spaces will have to integrate digital twin capabilities. They create a clear logical separation between information technology and operational technology as shown in figure 4:

- data is processed in the virtual entity of a digital twin,
- data is processed in the physical entity of a digital twin,
- data is exchanged between the virtual entity and the physical entity.

1.4.3 Trustworthiness support building blocks

As discussed above, the impact of the trustworthiness concern includes security, privacy, safety, reliability and resilience but also full stack integrity, The following processes and models can be considered:

- In order to characterise trustworthiness:
  - a trustworthiness characterization methodology to characterise (1) the attributes and (2) the context of use;
  - a attribute characterisation model (outcome of the methodology);
  - a context of use characterisation model (outcome of the methodology); and
  - a trustworthiness architecture representation model.
- In order to engineer trustworthiness:
  - an impact assessment process focusing on risk management activities;
  - a trustworthiness-by-design process to integrate trustworthiness considerations in the lifecycle process of a system; and
  - a system assurance process.
- in order to operate systems and maintain trustworthiness:
  - a system maturity process;
  - an integration process specifying how systems are integrated in a system of systems; and
  - a control model explaining how trustworthiness is controlled during operation.

These processes and models should lead to the provision of a trustworthiness support building block.

1.4.4 Interoperability support building blocks

Interoperability is defined as the ability for two or more systems or applications to exchange information and to mutually use the information that has been exchanged[20]. The support of interoperability in data spaces infrastructure integrating IoT and edge requires support along three dimensions.

The first dimension is the **interoperability facet dimension**. ISO/IEC 21823-1 defines five facets:(Interoperability for internet of things systems — Part 1: Framework)

- Transport interoperability (ISO/IEC 21823-2) which involves physical connections and signals, enabling data transfer between systems, using protocols of data transfer;
- Syntactic interoperability (ISO/IEC 21823-4) which involves data, enabling reception of data in an understood format, using standardised data exchange format;
- Semantic interoperability (ISO/IEC 21823-3) which involves ontologies, enabling reception of data using an understood data information model, using a common interpretation of data information model;
- Behavioural interoperability (on-going preliminary work item in ISO/IEC JTC1/SC41) which involves models, enabling the description of expected outcomes to interoperability operations; and
- Policy interoperability (on-going preliminary work item in ISO/IEC JTC1/SC41) which involves regulatory and organisational policies and interoperation context, using conditions and control of access and use.

The second dimension is the interoperability lifecycle and architecture, or **interoperability case dimension**:

- Each of the interoperability facets includes interoperability artefacts (physical connections and signals, data, ontologies, behavioural models, and policies) which must be engineered according to a lifecycle process. As described in the AIOTI position paper on semantic interoperability standard 21 , these lifecycle processes include a consensus stage where agreement is sought. An interoperability case is constructed and agreed upon by the stakeholders of the ecosystem having an interest to participate in this interoperability.
- A point of interoperability in the data space ecosystem is identified. Points of interoperability have an impact on the data space ecosystem architecture. The following architecture models can be involved:
    - Models to identify interoperability points in data spaces
    - Models to engineer interoperability cases for interoperability points in data spaces
    - Models to verify conformance of participating entities in data spaces, and
    - Models to maintain and update interoperability cases in data spaces.

[20] ISO/IEC 21823-1 (Interoperability for internet of things systems — Part 1: Framework)

[21] https://www.researchgate.net/publication/336677616_Towards_Semantic_Interoperability_Standards_based_on_Ontologies

Consequently, data spaces will have to integrate interoperability support building blocks, taking into account the interoperability facet dimension, as well as the interoperability lifecycle and interoperability architecture dimension.

A building block providing trustworthy data exchange is needed such as the data connector approach proposed by IDSA (see 2.2).

1.4.5 Infrastructure reconfiguration support building blocks

Some of the infrastructure elements in the IoT-Edge-Cloud continuum can cause changes to a data space ecosystem. These changes corresponds to lifecycle stages of elements:

- IoT devices added, changed, or removed
- Edge computing capabilities, added, changed or removed
- Cloud computing capabilities, added, changed or removed

These modifications can imply new emerging behaviours and emerging risks that must be taken into account in data spaces. Consequently, data spaces will have to integrate a building block to support reconfiguration of the data space infrastructure.

1.4.6 Data Business Marketplace Building blocks

The purpose of data spaces is to foster the use and exchange of data, while following the rules of data governance. To this end a data business marketplace building block is needed. It has to address the following data space challenges:

- federation of data spaces,
- usage control,
- support the other data space roles (data processing, data providers, data owners),
- support of the data lifecycle,
- support semantic and behavioural interoperability (Data Models),
- support the infrastructure continuum,
- data sovereignty
- trustworthiness, and
- governance.

The data business marketplace should typically integrate the above building blocks:

- data governance support,
- cyber physical system and digital twin support,
- trustworthiness support,
- policy support,
- Interoperability support, and
- infrastructure reconfiguration support.

1.4.7 Hyperdimensional Interoperability

Policy awareness and Context awareness

The vision of a smarter cyber-physical infrastructure requires full awareness of policy and context. AI systems for instance will require context-making to improve performance and explainability: they need to be policy-aware and context-aware. Data needs to be put into context for IoT and cyber physical systems to remain relevant and adaptive to changes in use cases and scenarios.

Context can be defined as follows:

- Context is multi-dimensional. As shown in Figure 12, it encompasses the following dimensions:
  - Semantic (meaning and logic)
  - Spatial (physical and situational)
  - Societal (values and value)
  - Systems (networks and ecosystems)
- As shown in Figure 11, Context is represented by meta-data models that describe the activities of people, places and things over time. Context needs to be shared between networks of heterogeneous devices and applications empowering them to proactively offer enriched, situation-aware and usable content, instructions and experiences.
- Context is made up of the elements of relationships between entities, objects, locations and actions—commonly known as the Who, What, When, Where, How and Why of any scenario, situation or circumstance. The answers to these questions are often stored in different data silos and different data spaces. They need to be made interoperable, shareable and addressable by multiple competing AI algorithms that can maintain their coherence at scale.

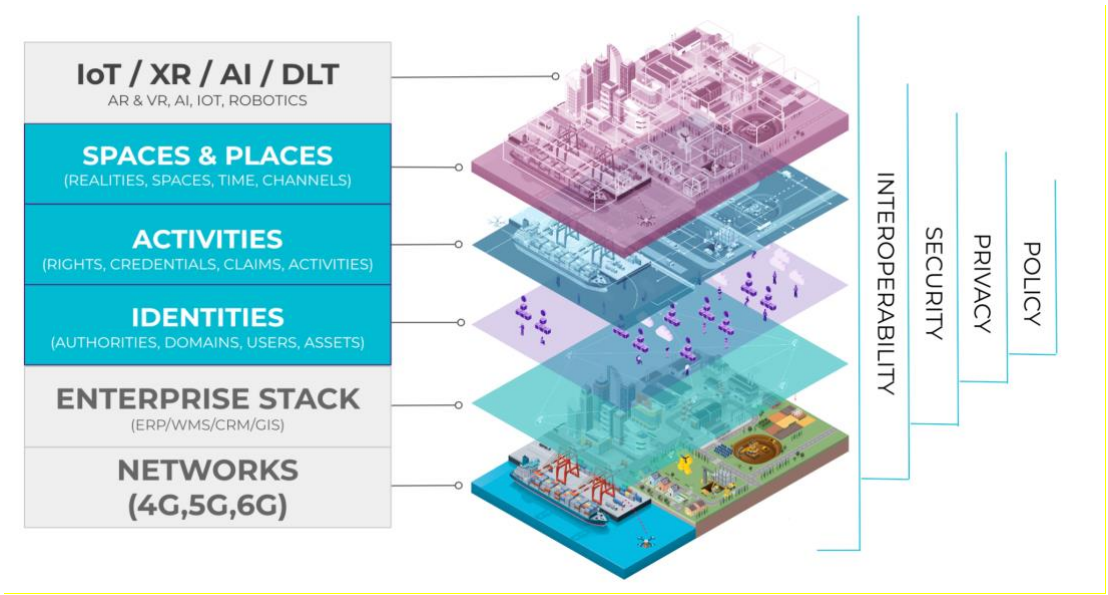


Figure 11 – Multi-dimension Interoperability

Spatial Web initiative and IEEE P2874 standards

The Spatial Web Foundation[22] has defined a contextual model and communication protocol that captures multi-dimension data interoperability. They are being standardised within the IEEE WG P2874[23], namely:

- HSML - Hyperspace Modelling Language
    - A common data model that enables adaptive intelligence at scale;
    - A standard that models spatial and hyperdimensional relationships which can exist between any base elements and their purpose.
- HSTP - Hyperspace Transaction Protocol
    - Multi-dimensional range query;
    - A contracting protocol that queries that language and sends and receives the common language's data.
    - Protocols allowing stakeholders to govern identities, activities and spaces and location in an interoperable way and across data domains.
    - Governs interactions between parties to ensure privacy and security.

Note that IEEE has designated HSML and HSTP as "Public Imperative". This designation is typically reserved for critical public infrastructure like nuclear energy, smart grids, and voting machines.

[22] https://spatialwebfoundation.org/

[23] https://sagroups.ieee.org/2874/

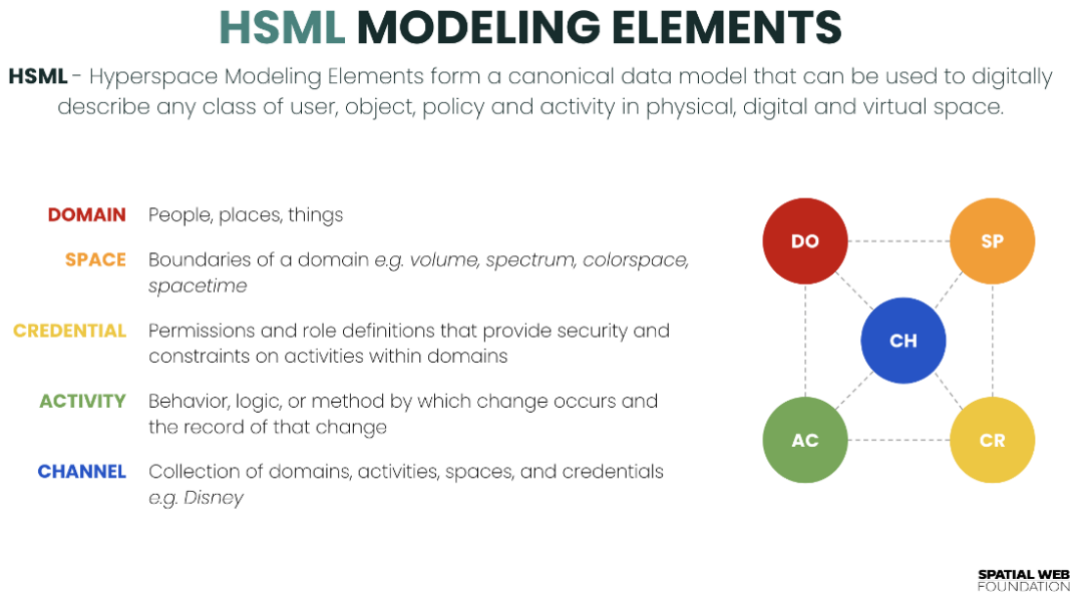


Figure 12 - HSML Modelling Elements, Source: Spatial Web Foundation

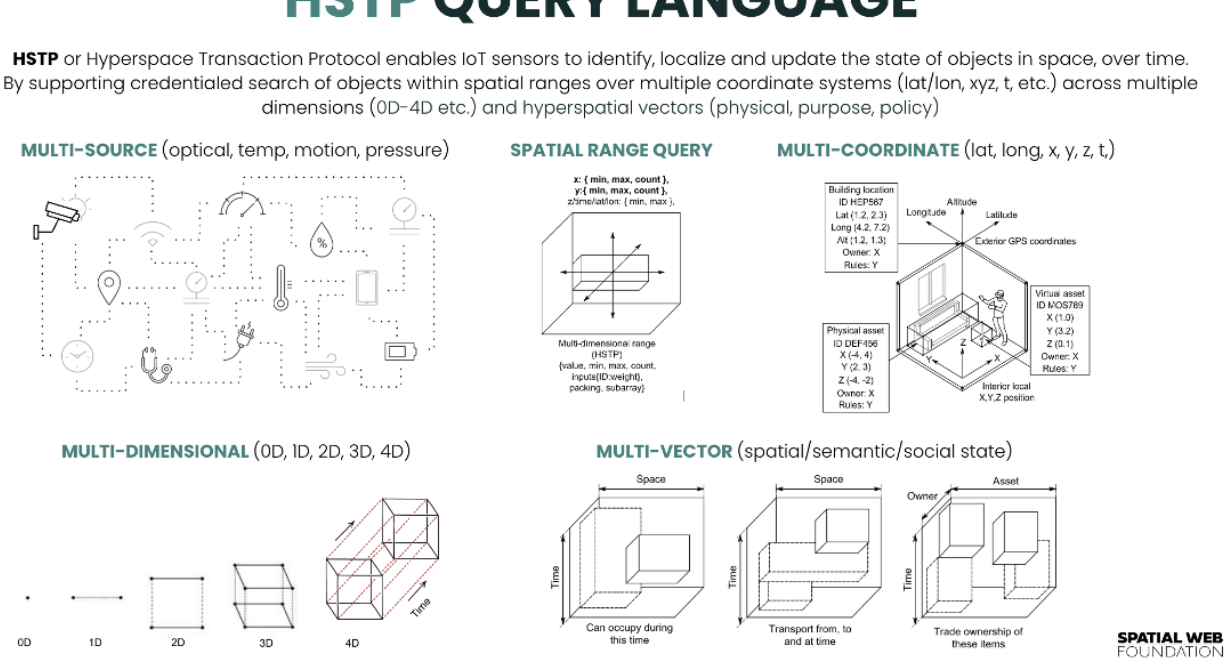


Figure 13 - HSTP Query Language, Source: Spatial Web Foundation

2 Relation to Solution Architectures

This section provides an analysis between the previous section which provided a reference architecture viewpoint on the integration of IoT and edge in data spaces. This section explains the relation with solutions architectures that are used in actual implementations:

- It explains how reference architectures specifications can be used to create a solution architecture
- It presents solutions architectures from international initiatives: IDSA, oneM2M, ETSI MEC.
- It presents a number of solutions architecture from European projects: Platoon, Interconnect, Smartbear, Assist-IoT.

2.1 Constructing solutions architectures

The approach to construct a solution architecture using reference architecture standards is based on the ISO standards guidance (based on ISO/IEC/IEEE 42010 architecture description, and JTC1/AG8 guidance work on meta reference architecture) as depicted in Figure 12:

- Reference architecture standards propose a number of common models as well as patterns which allow for customisation.
- A solution architecture is the result of grouping common models with the selected patterns.

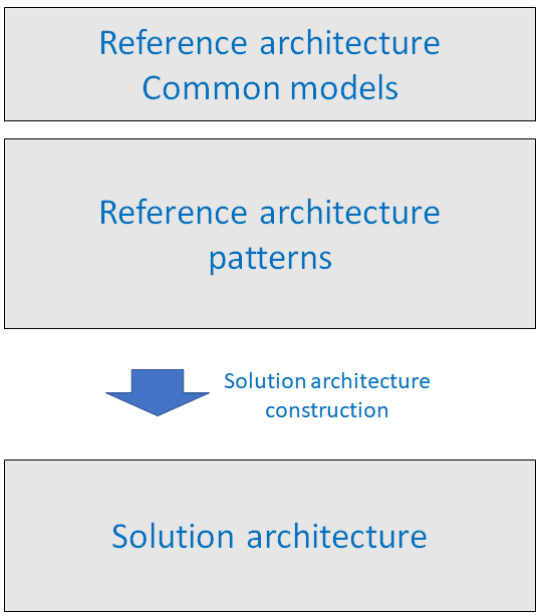


Figure 14 – Building a solution architecture

Figure 3 shows the result of creating a solution architecture for a data space ecosystem integrating IoT and edge: IoT and edge common models and data space common models are associated with patterns, IoT and edge patterns, data space patterns as well as domain patterns, for instance using SGAM[24] in the energy domain, or using RAMI[25] in the smart manufacturing domain.

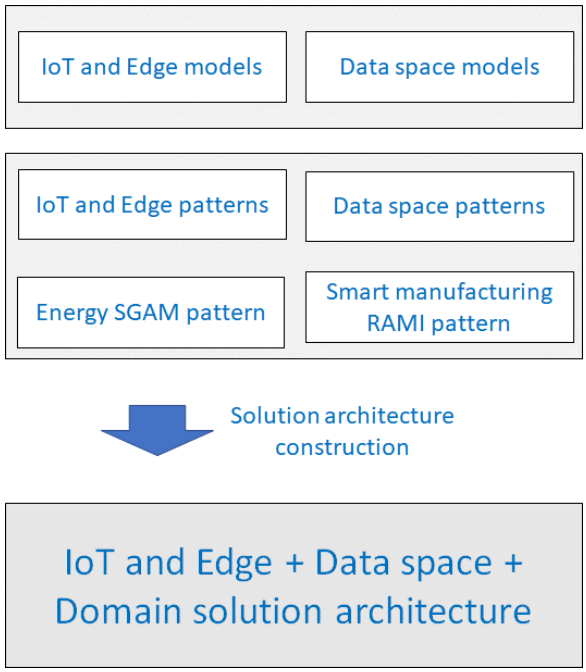


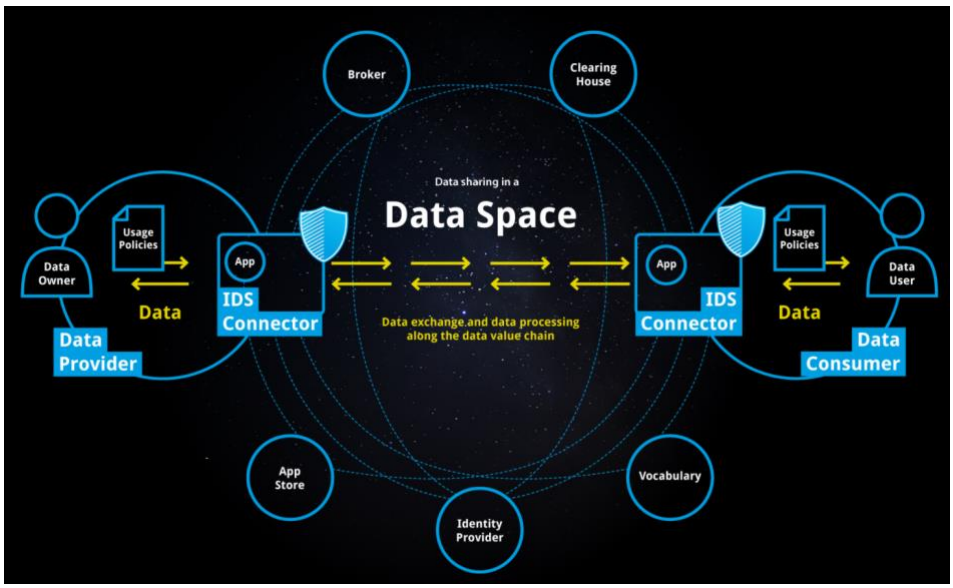
Figure 15 – Building a data space architecture integrating IoT and Edge

[24] https://ec.europa.eu/energy/sites/ener/files/documents/xpert_group1_reference_architecture.pdf
[25] https://www.beuth.de/en/technical-rule/din-spec-91345/250940128

2.2 IDSA Reference Architecture

2.2.1 Overall Characteristics

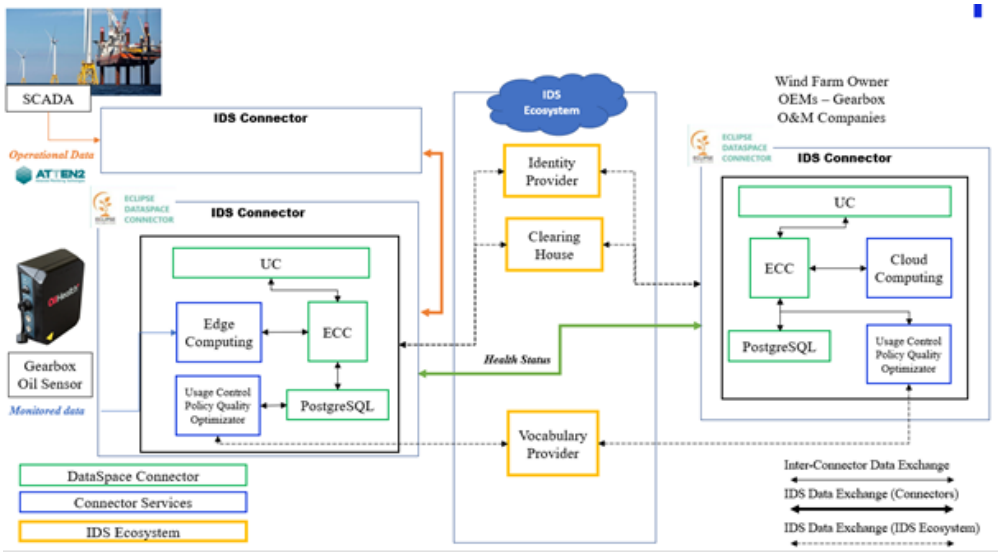
Table 8 – IDSA characteristics

Reference	https://www.internationaldataspaces.org/wp-content/uploads/2019/03/IDS-Reference-Architecture-Model-3.0.pdf (2019)
Description	The International Data Spaces Reference Architecture Model (IDS-RAM) sets the standard for building data-driven ecosystems, products and services that enable standardized, trustworthy and self-determined data exchange based on European values. The IDS-RAM upholds the data sovereignty of the creator of the data, guarantees trust among participants and ensures data privacy and security throughout the data exchange.
Stakeholders and concerns	Governance of the data spaces: Participants in IDS can assume different roles, the core participants are data owner, data consumer, data provider, data user or app provider. Intermediaries act as trusted entities that provide services, establishing trust, providing metadata, and creating a business model around their services. Another category are the IT companies providing software and/or services (e.g., based on a software-as-a-service model) to the participants of the International Data Spaces. The Governance bodies are the Certification Body, Evaluation Facilities, and the International Data Spaces Association. Other concerns are the data usage control, the usage contract enforcement, the data provenance and tracking and the new business models development for the data economy. Finally, the IDS Information Model addresses the main modeling concerns of data with the "C-hexagone": content, concept, community of trust, commodity, communication and context.
Architecture principles	The IDS standard enables trustworthy data exchange among certified data providers and recipients, based on mutually agreed rules. The IDS-RAM assures the data sovereignty for the creator of the data and secure data exchange and data processing along the data value chain based on European values, by which means equal opportunities through a federated design and decentralized architecture. The figure below illustrates the data exchange process between the core IDS-RAM components. The main component that makes this exchange possible is the IDS Connector, which is responsible for forwarding the data from the Data Provider to the Data Consumer.  **Figure 16 – IDSA Data space** Source: https://internationaldataspaces.org/why/data-spaces/

Data management capabilities	IDS-RAM provides various data management capabilities across its different layers: - First capability is to enable data sovereignty in data exchange and data sharing by providing a trust model for data spaces, data encryption, including identity management for organizations and components, and certification process (Conformity assessment programme). - IDS-RAM provides standards and technical means for Usage Control. Specify requirements (and conformity to those) for the trusted use of data across security domains. - In addition, the IDS Information Model provides the metadata interoperability expressed as an RDF ontology/metamodel for data as an asset. - Finally, the IDSA Rule Book provides a governance scheme for data spaces that describe the Business, Legal, Operational, functional and technical rules.
Roadmap	The IDS-RAM provides the possibility to implement processes regarding the following data management aspects, but they are not included as an integral part of the architecture: - Usage Control Enforcement: IDS-RAM cannot, and does not intend to, replace legal contracts or licensing agreements. Instead, the IDS provides a framework to technically enforce usage controls in addition to existing, legally binding contracts. Such enforcement would require the Connectors data flows to be modified to add monitoring and interception points (i.e., Policy Enforcement Points, PEPs) that request permission or denial of an action from a centralized decision engine (i.e., a Policy Decision Point, PDP). - Data Quality: can be assessed by extending the functionality of the Connector with self-implemented Data Apps that perform data quality checks before a data exchange is carried. - Data Provenance: can be controlled through local tracking components integrated into the Connectors and a centralized provenance component that analyses the logs from all data exchanges.

2.2.2 Integration of IoT and Edge Computing

Table 9 – Example of IoT and Edge Computing in IDSA use case

Connection to IoT and Edge (a.k.a. computing continuum support)	**TEKNIKER and ATTEN2 Wind turbines use case for data sharing in energy sector** The use case shows how the adoption of the IDSA Architecture boosts the exchange of the data monitored by the sensors installed in wind turbines, which belong to the wind farm owner and the OEM, with other interested third parties such as technology and component suppliers, as secure data exchange and data sovereignty is guaranteed.  **Figure 17 – Example of IoT and Edge integration in IDSA Data space** Developers partners: - ATTEN2 Condition monitoring solutions provider (https://atten2.com) - Tekniker IDSA compliant developer (https://www.tekniker.es)
Data governance across computing continuum	**ONCITY use case from German Edge Cloud** ONCITE is a plug & produce solution in the form of a compact computing centre that is based on highly available and scalable edge cloud technology. The solution allows suppliers and manufacturers to make their data available for being used in digital processes (e.g., industrial AI or industrial track & trace) in real time and at the highest possible level of security. With ONCITE, companies can process and store data on site, before they exchange it over a public cloud – with data sovereignty being ensured for each data owner/supplier across the entire process.

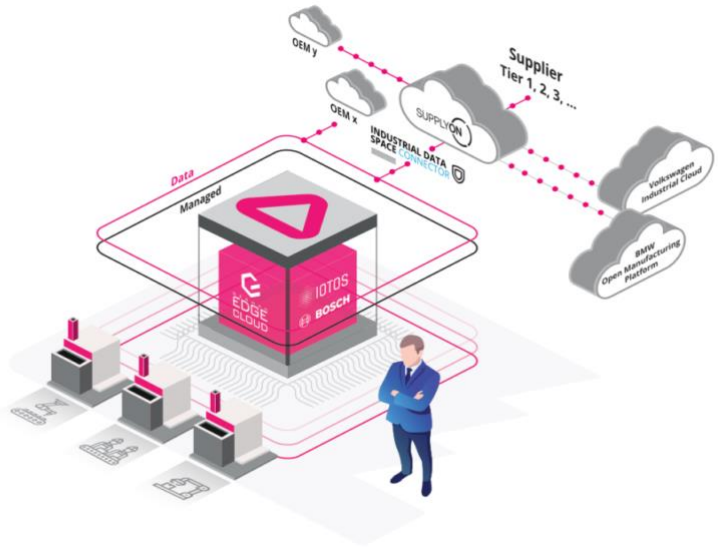


Figure 18 – ONCITY Data governance

Source: https://internationaldataspaces.org/usecases/german-edge-cloud/

Interoperability

GAIAboX- Secure resource management, file storage and data exchange in IDS by Nicos AG.

The idea of GAIAboX is prototyping a secure and sovereign resource management and file storage. It is accessible via FTP, SSH and HTTP and follows the concept of "Linked Data Platform" (LDP, see W3C). The achievement of GAIAboX is not limited to data sharing, as open inventory platform it also allows representing any resource. GAIAboX will also provide application protocols like mqtt, gRPC, WebSockets in order to make "publish/subscribe" available. It intends to provide data and information in a standardized and semantically described manner according to the concept of the Asset Administration Shell (AAS), thus enabling interoperability and easy interaction.

Maritim data space by SINTEF and DNV GL

Maritime shipping companies are required by law to transmit a set of important data before entering every port. Providers and organizations on shore do not have access to ship data like emissions, fuel consumption and route details. Data access agreements have to be negotiated individually. A common ecosystem for data exchange from ship to shore that simplifies the process is needed.

The IDS-based Maritime Data Space brings together all participants and platforms in one trusted and secure data ecosystem.

Source: https://internationaldataspaces.org/usecases/sintef-maritime-data-space/

2.3 oneM2M

2.3.1 Overall Characteristics

Table 10 – OneM2M characteristics

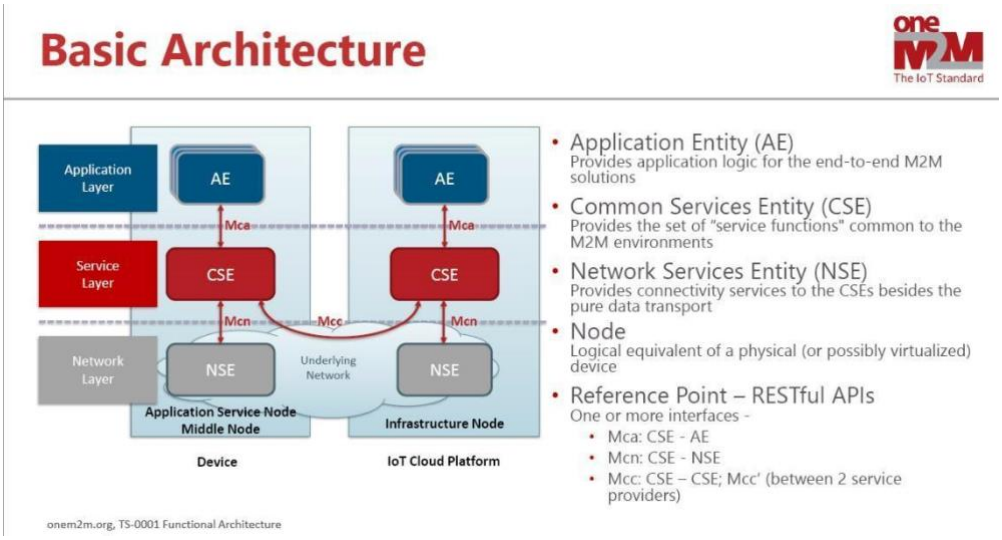
Reference	https://www.onem2m.org/using-onem2m/what-is-onem2m
Description	oneM2M is a global partnership project founded in 2012 and constituted by 8 of the world's leading ICT SDOs. The goal of the organization is to create a global technical standard for interoperability concerning the architecture, API specifications, security and enrolment solutions for M2M and IoT technologies based on requirements contributed by its members.
Stakeholders and concerns	More than 200 members of 8 of the world's leading ICT standards development organizations, notably: ARIB (Japan), ATIS (United States), CCSA (China), ETSI (Europe), TIA (USA), TSDSI (India), TTA (Korea) and TTC (Japan) from many diverse business domains.
Architecture principles	oneM2M is based on a RESTful architecture. Its interworking framework can simultaneously interwork different IoT device technologies with one another and provides an abstract and simplified API for applications to communicate with devices. oneM2M Service Layer supports configurable access policies that define clear rules dictating, for each resource, who is authorized to access and what operations are allowed under which conditions. 
Data management capabilities	- oneM2M interworks different IoT device data models with one another (e.g., OCF, LWM2M). all devices are presented to the App via oneM2M API. Via standardized oneM2M API, App developers can use device services and manage devices. - Once the data model is abstracted into oneM2M, App developers can access all devices in a common manner and make use of oneM2M value-add capabilities such as resource discovery, generating events via subscriptions and notifications, grouping and access control.
Roadmap	Release 5: - oneM2M system enhancements to support data protection regulations - Effective IoT Communication to protect 3GPP networks - oneM2M and sensor things API - Advanced semantic discovery - System enhancements to support data license management

Figure 19 – oneM2M basic architecture

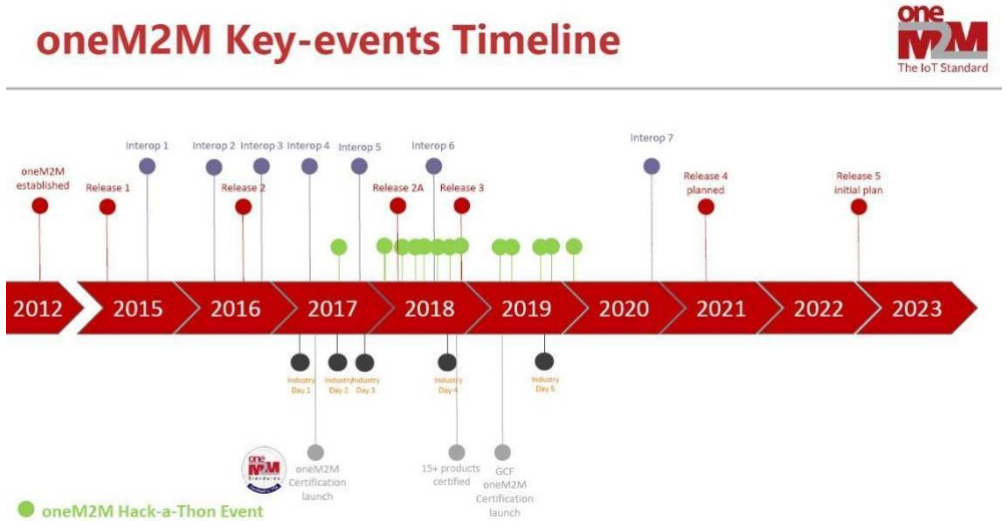


Figure 20 – oneM2M Roadmap

2.4 ETSI (Multi-access Edge computing)

2.4.1 Overall characteristics

Table 11 – ETSI MEC characteristics

Reference	https://www.etsi.org/technologies/multi-access-edge-computing
Description	The Multi-access Edge Computing (MEC) initiative is an Industry Specification Group (ISG) within ETSI. The purpose of the ISG is to create a standardized, open environment which will allow the efficient and seamless integration of applications from vendors, service providers, and third parties across multi-vendor Multi-access Edge Computing platforms.
Stakeholders and concerns	A continuously growing membership (now at 125 between members and participants), across all the stakeholder categories of the whole value chain (from operators, technology providers, research institutions, public administrations, SMEs, startups, …). Also, GSMA and 5GAA (5G Automotive Association) have joined MEC, and the ISG is establishing collaboration with many groups and open-source communities (e.g. LF, Akraino).
Architecture principles	Multi-access Edge Computing (MEC) offers application developers and content providers cloud-computing capabilities and an IT service environment at the edge of the network. This environment is characterized by ultra-low latency and high bandwidth as well as real-time access to radio network information that can be leveraged by applications. ETSI MEC architecture support multiple access technologies, such as 5G, Wi-Fi and fixed networks. It exposes a set of RESTful APIs to edge applications, to support multiple use cases, including MEC V2X API and MEC IoT API (which can interoperate different IoT devices and IoT service platforms).

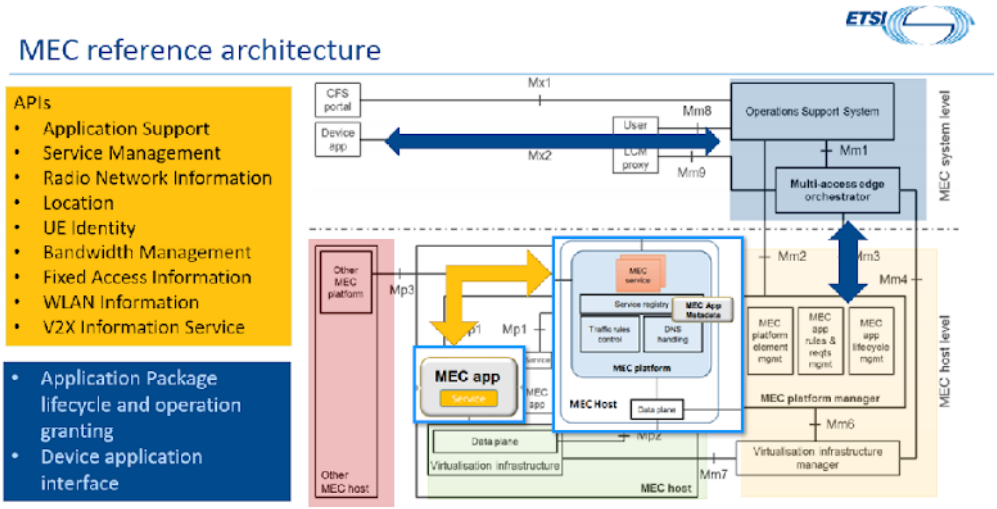


Figure 21 – ETSI MEC Reference architecture

Data management capabilities	- MEC Applications can consume and/or produce MEC service APIs. Also, New APIs (compliant with the MEC API principles) can be added and exposed in the MEC Platform. MEC also enables applications and services to be hosted 'on top' of the mobile network elements, i.e., above the network layer. These applications and services can benefit from being in close proximity to the customer and from receiving local radio-network contextual information. - MEC architecture variant for MEC Federation introduces the concept of a "federated model of MEC systems enabling shared usage of MEC services and applications". This MEC Federation concept is a key enabler for supporting the requirements coming from GSMA OPG (Operator Platform Group): it enables inter-MEC system communication and allows 5G operators to collaborate among them and with service cloud providers and other stakeholders. - The ISG also actively works to help enable and promote the MEC ecosystem by publishing in the ETSI forge website [26] the OpenAPI representations of published MEC service APIs.
Roadmap	MEC Phase 3 (ongoing) includes: - Completion of outstanding Phase 2 work - MEC as heterogeneous clouds: Expanding traditional cloud and NFV LCM approaches; Inter-MEC systems and MEC-Cloud systems coordination: "MEC Federation" (MEC 035 – published / MEC040 – ongoing); Mobile or intermittently connected components, and resource constrained devices (MEC 036) - MEC Security (GR MEC 041) - MEC deployments: MEC in Park enterprises (MEC 038) - Continuing emphasis on enabling edge application developers: - Application Package Format and Descriptor Specification (MEC 037) - API Serialization (and maintain completed APIs) - Sandbox development - Testing and compliance

[26] https://forge.etsi.org/rep/mec

2.4.2 Integration of IoT and Edge Computing

Table 12 – Integration of IoT and Edge computing in ETSI MEC

Connection to IoT and Edge (a.k.a. computing continuum support)	Multiple use cases and requirements in MEC are related to IoT space (MEC 002, published in 2022). The MEC IoT API introduces a MEC service to assist the deployment and usage of devices that require additional support in a MEC environment, e.g., due to security constraints, limited power, compute and communication capabilities, such as IoT and MTC devices. The work item is defining an API necessary to enable the device provisioning, and configuration of the associated components and applications requiring connection to these devices. Furthermore, an ongoing work item (MEC 036) studies MEC in resource constrained terminals (fixed or mobile), and in particular how terminal units, mobile hosts and personal devices can be used to support cloud computing at the edge. The study will focus on these aspects: - Limited availability of compute resources for running MEC applications and it's impact on life cycle management of VMs, Containers or other form of virtual instances. - Mobility of constrained terminals impacting reachability of MEC applications, maintenance of reasonable connectivity, device availability and discovery of appropriate services. - Impact of unavailability of reliable high bandwidth backhaul connectivity (e.g., wired or wireless). - Security and authorization to use a constrained terminal, privacy of user data Applicability of MEC Specification to support cloud computing on such constrained environment will be studied
Cyber physical systems and digital twins	Digital Twins cover a vast domain of applications and use cases, that may need MEC support in various vertical market segments. Few examples relevant for MEC APIs: - MEC V2X API – supports predicted QoS information exposed at application level, e.g. helping automotive use cases on automated and connected vehicles - MEC IoT API –provides means to incorporate heterogeneous IoT frameworks in MEC, and exposes APIs for the MEC platform configuration to facilitate the device provisioning and the configuration of the IoT components running as MEC applications. Moreover, a new study MEC 044 on "Abstracted Radio Network Information for Industries" is covering different needs for the abstracted information and a few different industry areas e.g., AR&VR, V2X, Logistics, Future Factories, Coordinated Robots and Drones. The aim for the abstraction is to provide a developer-friendly API that hides the complexity and requires only little technical skills or knowledge of the underlying Radio Network.
Data governance across computing continuum	As MEC Infrastructures can span a wide geographical distribution and be located in challenging environments, maintaining a uniform data-centre level of physical security in data governance is a significant challenge. In that perspective, MEC systems shall comply with regulatory requirements for lawful interception and retained data (ref. ETSI TS 101 331 and ETSI TS 102 656). Moreover, the confidentiality and data integrity of all messages should be ensured by using TLS on each interface of the MEC Architecture. Appropriate security controls are required for protecting sensitive data storage, processing, and transfer by MEC applications. The MEC platform should authenticate all MEC application instances and only provide them with the information for which the application is authorized. OAuth 2.0 based on X.509 client certificates are used for authorization of access to RESTful MEC service APIs defined by ETSI ISG MEC. In case of service-producing applications defined by third parties, other mechanisms such as standalone use of JWT can be used to secure related APIs.

Trustworthiness across computing continuum	The MEC system shall provide a secure environment for running services for the following actors: the user, the network operator, the third-party application provider, the application developer, the content provider, and the platform vendor. Furthermore, the MEC platform shall only provide a MEC application with the information for which the application is authorized. On LI&RD requirements (as identified by the NGMN MEC Security report), there is also a special requirement to ensure target information is protected appropriately, potentially with a hardware root of trust, or utilizing a dynamic triggering model that minimizes the sensitive information available on the MEC element. The current "MEC Security" work item (MEC 041) will study security topics and paradigms that apply to MEC deployments. The study will broadly cover the themes of application and platform security, Zero-Trust Networking, and security requirements for MEC Federations. It may also draw upon prior work from other standards and gather requirements from industry associations (e.g., 5GAA). It will identify gaps in ETSI ISG MEC and provide recommendations for new normative work.
Interoperability	Interoperability is a key aspect, including many areas where MEC is critically needed: - MEC Federation: multi-operator environments are key scenarios for automotive use cases (e.g., as required by 5GAA), and the standard support of GSMA OPG requirements is critical for the interoperability in this heterogeneous scenario. ETSI MEC is working on Federation enablement APIs (MEC 040), - Standardized APIs (or new APIs compliant with the MEC API design principles in MEC 009) are also a key enabler for interoperability across multiple stakeholders - More in general, the MEC harmonized architecture with 3GPP EDGEAPP (ref. ETSI white paper n.36) is key to offer an interoperable environment, where consistent standards can open the edge computing market, avoiding duplication and market fragmentation.

2.5 Flying Forward 2020 research project and the Spatial Web architecture

2.5.1 Overall Characteristics

Table 13 – Project characteristics

Reference	Flying Forward 2020: supporting the future of UAM with COSM's geospatial governance capabilities and sensor fusion www.ff2020.eu
Description	Flying Forward is a 3-year collaborative research project to develop a new Urban Air Mobility ecosystem by incorporating UAM within a geospatial data infrastructure of cities. Hyperdimensional interoperability (see 1.4.7) is at the heart of that new infrastructure. COSM[27] is a example of hyperdimensional interoperability solution which uses sensor fusion to provide the capabilities to govern the behaviour of autonomous drones flying in urban areas. That involves integrating legal and spatial policies in the simulation and activation of the scenarios, while for example allowing for dynamic (re)routing, multi-party interoperability and managing landing zones. The FF2020 demonstrators cover 5 use cases : - Security and surveillance - Infrastructure inspection - Deliveries - Crowd management - Emergency deliveries.
Stakeholders and concerns	Enterprise operators (logistics, mobility, healthcare, agriculture, construction, energy, manufacturing, etc) Governments and regulators concerned with the digital policies managing autonomous and AI systems. End users and citizens who are concerned with privacy, security, prosperity and digital policies that benefit all of humanity.
Architecture principles	The project leverages the spatial web protocols to define the geospatial digital infrastructure to be able to govern the safe behaviour of autonomous drones. More specifically, COSM provides an Artificial Intelligence context aware platform built upon the open protocols HSML and HSTP for managing resources and deploying AI applications on the Spatial Web. Its five Flow Modules are designed to address the requirements for a universal network of humans, machines, and AI (see roadmap section below for more details). In these scenarios, the geospatial digital infrastructure leverages spatial twins of each living lab location (augmented digital twins of campuses, hospitals, city centres…) and goes beyond Device Identification and Profile Authentication to make devices become Policy-aware and Location-aware. As a result, physical activities of IoT devices can be governed and enforceable at the hardware level. COSM allows interoperability between the edges and systems in involved in the use cases and in particular allows the behaviour of any "thing" (drone, car, truck, vessel…) to be law abiding and adaptive in real-time to policy changes and updates

[27] https://www.verses.io/cosm-os

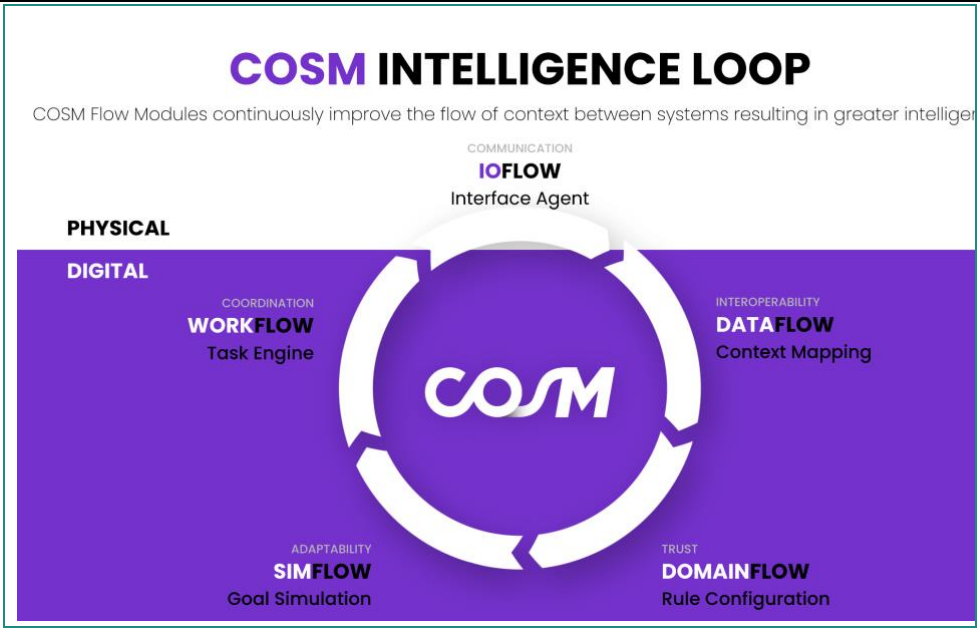


Figure 19 – COSM

Data management capabilities

Below is a description of the 5 data management modules provided by COSM:

- IO-Flow (Communication): Interface agent. Management of connected devices to publish sensor data and instruct machines and humans to perform tasks;
- Data-Flow (Interoperability): Context mapping. Normalization and correlation of data from disparate systems into a unified coherent HSML Context Graph™;
- Domain-Flow (Trust): Rule configuration. Definition of the policies, permissions, and credentials that govern the interactions of all actors on the Spatial Web;
- Sim-Flow (Adaptability): Goal simulation. Simulation of goal-based optimizations based on requirements or restrictions, and recommendation of tasks to achieve ideal states;
- Work-Flow (Execution): Task engine. Delegation of optimization tasks to a connected workforce of humans, autonomous vehicles, and/or bots.
- Spatial DNS: In order for actors, autonomous vehicles and bots to operate and for AI to orchestrate their activities, a query method is required to gather all the data necessary such as:
    - Where can I go?
    - What credentials do I need?
    - How should I go there?
    - What can I do?``

Roadmap

To validate the FF2020 solutions and approach, tests will be conducted in 10 demonstrators in collaboration with 5 Living Labs across Europe (Eindhoven, Milan, Tartu, Oulu, Zaragoza). The Living Lab partners will provide high-impact demonstrators, and specify their needs and requirements to ensure compliance with regulations, in terms of non-functional properties, safety, tools and processes.

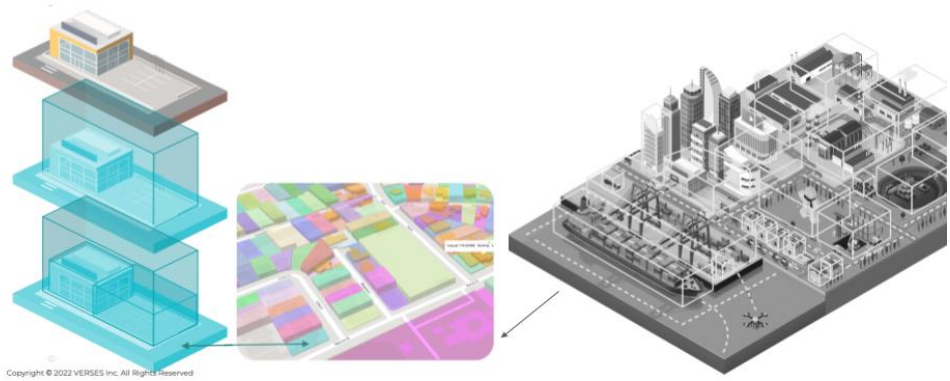
The consortium has successfully delivered the Eindhoven demonstrator for security, infrastructure maintenance and delivery. Zaragoza is scheduled for october'22 and Milan, Tartu and Oulu for 2023.

We have defined three stages of development in order to reach our full potential. The three stages are:

- Stage 1: Regulatory, governance and technology R&D as well as technical integration
- Stage 2: Experimentation and validation
- Stage 3: Regulatory, governance and technology integration to propel scalability and sustainability (with EASA regulation in particular)

2.5.2 Integration of IoT and Edge Computing

Table 14 – Use cases using COSM

Connection to IoT and Edge (a.k.a. computing continuum support)	There are use cases where this is needed in the areas of logistics, spatial analytics and mobility. This is supported by the following features: registering and connecting drones, bots, handhelds, HMDs, IoT, etc that are spatially routed using a local 3D model of the location to complete tasks in the most efficient manner possible. - At each major node of the system such as a building, intersection or city, a COSM instance will reside on local edge computing hardware linked by a robust configuration of WiFi, 5G and and Bluetooth networks. The local node will store up to date information about the 3D model, rules, fees, other actors in the space, their priority and their path. To ensure the least latency and uptime, all the major algorithms and compute power will take place at the edge.
Cyberphysical systems and digital twins	The geospatial infrastructure leverages "spatial twins" – digital twins of locations augmented with specific data streams and Spatial Domains. The latter are digital titles linked to 3D volumetric locations such as buildings, ports, streets, or larger regions, such as cities, states, continents and trading blocs. Spatial subdomains represent subspaces that have a holonic structure, which allows for the orchestration of hierarchical rights and policies.  Figure 20 – Spatial domains Spatial domains enable secure management of digitally mediated rights and permissions for: - Who/what is authorized to access the domain; - What content or data is available to view; - Who can publish and modify content; Who can transact or interact with it. They contain all the rules, rights, permissions and fees associated with a geospatial region as defined by their owner or authority. This results in a governance layer of geospatial information written in HSML that can be queried by actors as they approach and move through them. The 3D model contains all the geometries, addresses, spatial anchors, location of IoT devices, and any subdomains. The 3D model is dynamic and can be updated in real time as local conditions change. A visual digital twin can be used to represent the spatial domain and the activities happening within it. Routing data defines the designated and available routes through a space, the quickest path between any two locations and any speed or occupancy limits. Contracts are applied to a space or asset to perform certain activities such as moving an asset from one space to another for a fee. Smart wallets can be held by any actor, space or asset that is performing activities or holding a contract. A smart wallet contains the profile, financial accounts, credentials, etc and can autonomously complete transactions.
Data governance across computing continuum	Data governance is needed for sharing anonymized datasets with AI privately and securely. The use of unique decentralized identifiers (DIDs) for each interaction between any identity and any service on the network. DIDs can either be public, which is required for any Authority and expressed as JSON-LD documents, or private for each pairwise session between two parties. No unique data about an individual is ever stored on third party systems that can be used to correlate across services without their consent, yet data can still be aggregated for federated learning protocols through the use of zero-knowledge verifiable credentials. Each HSTP packet on the network is defined in HSML as a JSON-LD document where data access activities are also transparently expressed in the HSML context elements. All records related to individuals are stored privately by default, and are only able to be aggregated when cryptographically proven to be uncorrelatable across population sets with

	clear transparent consent from the individual to access records and their express right to be notified.
Trustworthiness across computing continuum	Trustworthiness is needed for sharing anonymized datasets with AI privately and securely. Data are persistent and can be accessible according to terms expressed in the DID documents associated with schemas, which can then be audited based on the transparent expression of rules around control of records described in the DID documents themselves. It can be supported as follows: - End-to-end digital and physical asset tracking Both physical and digital assets are tracked from origin where the digital twin is updated and enriched with new context, stored in either a public DID and/or private DIDs. - Transparent chain of custody and independently auditable transactions Each actor that interacts with any asset does so in a permissioned way, where all transactions by default are designed to be auditable without counterparty risk of each party participating in the transaction. - Real-time contract validation and enforcement Contracts can be evaluated in real-time and at the edge of the network, which both improves performance and increases transparency such that client and server, edge and cloud, can have a priori parity on the contracts and code execution between them. - Chained contracts across locations and identities Workflows can be consistently designed across different permissioned identities, assets, and spaces, and then chained together so that additional trust gets built up with every interaction.
Interoperability	Interoperability is needed in the areas of edge devices (IoT, mobility, CV, handhelds) and AI/ML applications working together or competing to determine the optimal outcome. Data schemas and fields are networked and interconnected with all others, making it possible to find reusable pathways to connect and convert between data formats and field types. The following has to be supported: - Globally registered schemas and schema version changes Any schema used to create data, issue or validate credentials, is cryptographically registered on the network to make schema changes more coordinated and prevent breaking changes across systems. - Registration of all services, endpoints, protocols in DID documents All mechanisms of interacting with any service connected to an identity are associated with the asset and updated in the global trust graph as new entries are added. - Reusable integration pathways between schemas and data types Each time an integration is done between two systems, whether converting data types or semantic synonyms between fields, this pathway becomes registered to the network as a code pipeline and can be intrinsically reused across contexts. - Contextually reconfigurable contracts and code Every contract expresses a unique context, but done in generic primitives designed such that they can be transferred across contexts such as a new location and be able to understand how to adapt based on spatial, semantic, and social anchors.

2.6 PLATOON IoT research project

2.6.1 Overall characteristics

Table 15 – PLATOON characteristics

Reference	PLATOON (Digital PLAtform and analytic TOOls for eNergy), https://platoon-project.eu/ Information kindly provided by Erik Maqueda (Technical manager, PLATOON project).
Description	PLATOON aims to develop a federated platform for the energy sector focusing on the following pillars: interoperability, trust and data analytic services. The project will develop , implement and validate into seven large scale pilots scalable and replicable solutions that accelerate energy transition.
Stakeholders and concerns	Consortium partners cover the whole energy value chain: large companies and SMEs (Distribution system operators (DSOs), Energy service companies (ESCOs), large Energy Consumers, ICT Companies…), Public Administrations, Research and Technology Operators (RTOs) and Academia. They have the following main concerns: - central platforms, vendor locking and interoperability issues - data privacy, security and sovereignty, and - lack of digital skills and tools to extract whole value of energy data.
Architecture principles	PLATOON has developed an open-source reference architecture based on widespread open reference architectures such as FIWARE, IDSA, COSMAG and SGAM.
Data management capabilities	The following features are supported: - Market place with IDS capabilities (Broker + App Store + Clearing House) - Access control and Authentication, using IDS data access protocols - Federation: defined reference architecture (see above), defined common semantic data models based on standards (SAREF, CIM, SEAS, OntoWind) - Trustworthiness management with the development of an IDS open source connector (TRUE connector) - Usage control: Data Usage Data App compatible with IDS connector. - Privacy compliance: Data Privacy Data App (CAPE) compatible with IDS connector.
Roadmap	- 1st version of open source IDS Connector with data usage and privacy features already available in PLATOON GitHub repository. - 1st version of open source Broker already available in PLATOON GitHub repository. - Rest of features should be ready by end of 2021. - In 2022 these features will be validated in large scale pilots.

2.6.2 Integration of IoT and Edge Computing

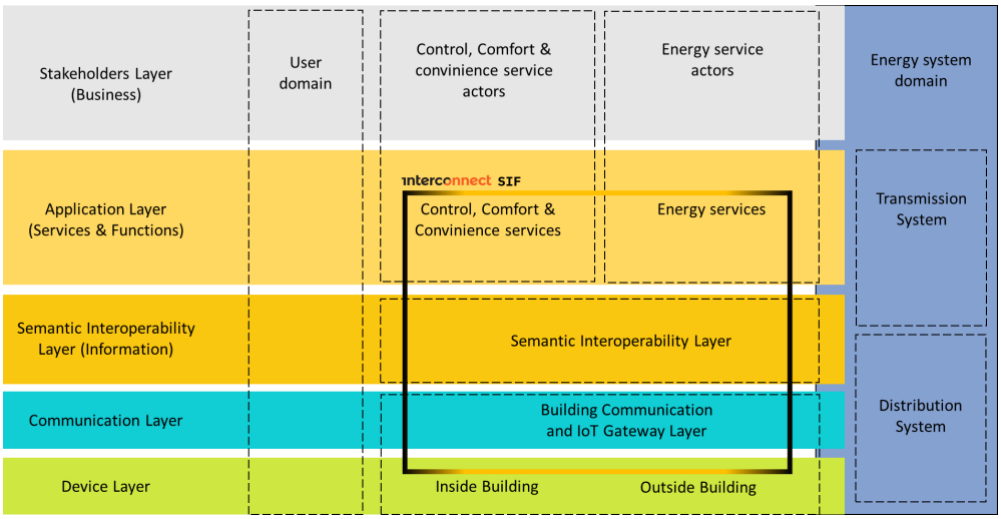
Table 16 – Integration of IoT and Edge computing in PLATOON

Connection to IoT and Edge (a.k.a computing continuum support)	PLATOON includes 6 large pilots that will develop, implement and validate different tools including data analytics at the edge. Computing continuum support is provided as follows: The Open-Source PLATOON Cloud-Edge framework will be ready by end of 2021 and openly available in the PLATOON GitHub repository. This framework is formed of several open source tools that cover the following features: - Node Management: Node status monitoring and visualization. - Analytics deployment in nodes.
Cyber physical systems and digital twins	One PLATOON large scale pilot involves cyberphysical systems and digital twins. It consists of an hybrid digital twin (physics based enhanced by data driven Machine Learning techniques) for Wind Turbines power drive train (Electric Transformer + Power converter). A first version of the hybrid digital twin should be ready by end of 2021. The solution is proprietary but we are currently working on a paper to disseminate results.
Data governance across computing continuum **Trustworthiness across computing continuum**	There are two 2 large scale pilots where data governance and trustworthiness in the computing continuum are important. Platoon is working on a Federated Edge Platform with an implementation of an open source IDS connector to be used at the edge to be able to send high frequency data. A mock-up is ready and a POC should be ready by April 2022.
Interoperability	Interoperability is of the main pillars of the project. PLATOON has defined an open common semantic data models based on standards (SAREF, CIM, SEAS, OntoWind). A first version of open data models is ready. It should be uploaded to the Github repository soon.

2.7 INTERCONNECT IoT research project

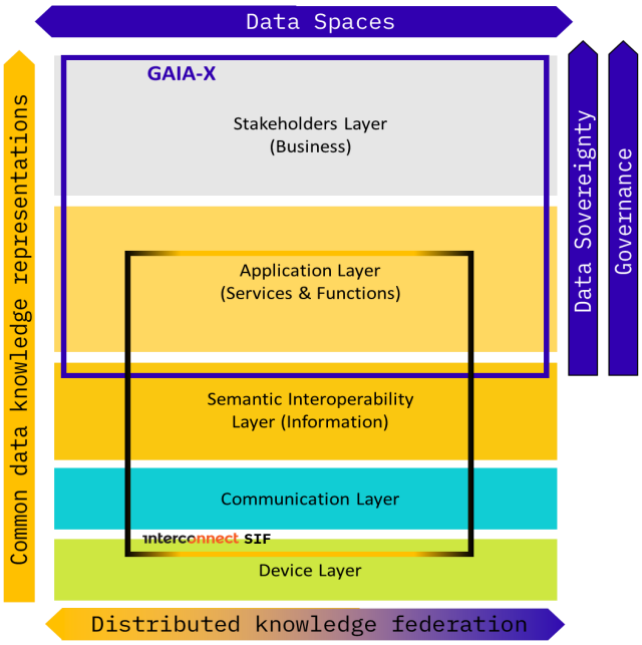
2.7.1 Overall characteristics

Table 17 – INTERCONNECT characteristics

Reference	https://interconnectproject.eu/
Description	Interoperable solutions/services connecting (devices in) Smart Homes, Buildings and Grids for the democratization of efficient energy management, through a flexible and interoperable ecosystem where demand side flexibility can be soundly integrated with effective benefits to end-users.
Stakeholders and concerns	An ecosystem of stakeholders with a need to interconnect devices and services: owners, facility managers and inhabitants of buildings, device manufacturers, IoT platform providers, energy service providers, energy (i.e., electricity) providers/retailers, Distribution System Operators (DSO), etc.
Architecture principles	Multiple architectural viewpoints (i.e., Energy, IoT, Interoperability Framework and Semantic) provide bridges for collaboration by domain experts versed in existing reference architectures (SGAM, RAMI, AIOTI, etc.). All viewpoints are based on separation of concerns and abstraction of functionality.  **Figure 22 – INTERCONNECT architecture**
Data management capabilities	Each InterConnect devices and/or service has a so-called Knowledge Base (KBs) associated with it. Knowledge Graphs (KG) are used for the exchange of Knowledge between these KBs. KGs are encoded using semantic web technology. Specific ontologies have been defined by the InterConnect project to have a shared understanding of the many concepts in the InterConnect ecosystem. The basic principle underlying the exchange of knowledge is to 'share on a need-to-know basis'. Each KB determines if it wants to share knowledge based on the content of received knowledge graphs. There is a Service Store where services can register (as a KB) and where they can be discovered by other KBs for the purpose of the exchange of knowledge. As InterConnect is a layer 'on top of' web/internet technology layers for the semantic exchange of information, it can make use of underlying functionality regarding access control, authentication, etc. This layer is also known as the Semantic Interoperability Layer (SIL).
Roadmap	The specific InterConnect technology that enables device manufactures, IoT platform providers, (energy) service providers and DSOs to interconnect using the Semantic Interoperability Layer is being finished to be used in 7 large scale pilots across Europe right now. Pilots should be running in 2022. It is proposed to investigate if it is possible to include finer-grained access control at the semantic level into Knowledge Graphs.

2.7.2 Integration of IoT and Edge Computing

Table 18 – Integration of IoT and Edge computing in INTERCONNECT

Connection to IoT and Edge (a.k.a computing continuum support)	All InterConnect use cases / pilots need the interconnection of devices in homes and buildings to both the grid and services in the cloud. Using the web/internet functionality InterConnect is location transparent with respect to Knowledge Bases. They could be located near IoT devices or at the edge (in/near homes and buildings), but they can also be in the cloud. In that case IoT platforms take care of transporting of relevant data to these cloud-based Knowledge Bases. At the level of the Semantic Interoperability Layer there are Knowledge Bases, no physical locations. The picture below shows how Interconnect plans to be integrated in future data space initiatives.  **Figure 23 – INTERCONNECT integration with GAIA-X**
Cyberphysical systems and digital twins	There are pilots where it is important to have knowledge available (at a coordinating party like an Energy Management Service provider) on the current and future (planned) production/consumption of electricity by devices in homes and buildings, to be able to avoid congestion and loss of production/consumption balance on the electrical grid. The InterConnect ontologies include semantic concepts that enable (IoT) device manufactures to express current and future use of electricity/energy. Several 'control model' semantic concepts are available, enabling flexibility service providers to calculate electricity production (e.g., in case of batteries) and consumption. A device Knowledge Base can share this 'control model' through a knowledge graph with the Knowledge Base of an energy management service provider.
Data governance across computing continuum	The services offered in InterConnect use cases often require knowledge on (the usage of) devices, it must be sent to service providers for (indirect incentivized or direct) remote operation of these devices for flexibility services. If not, energy management service providers cannot perform the automatic management of these devices based on the preferences of device users/owners. By restricting the amount data that is put in a Knowledge Base a device owner/manufacturer can control what is shared with the InterConnect Ecosystem. By not giving access to certain other Knowledge Bases device owners/manufacturers and service providers can restrict the spread of data/knowledge between different parties in the ecosystem.
Trustworthiness across computing continuum	Use Cases where a remote service in the cloud can switch on/off devices (consuming relatively large amounts of power) require a large amount of trust.

	The ecosystem of InterConnect can make use of the underlying functionality of the web/internet. On the top level of InterConnect's Semantic Interoperability Layer, the Knowledge Bases can decide which other Knowledge Bases are allowed to exchange knowledge. No trust means no acceptance. Also, a Knowledge Base can always decide not to update knowledge according to an incoming knowledge graph, if the new knowledge would result in an unstable or dangerous situation (e.g., for a device). To that, there is also the classical physical safety net for the flow of electricity (fuses, smart meter, etc.). It is also investigated if it should already be possible to add a trust indicator to devices and/or services. Currently the focus is on finishing a first version of the Semantic Interoperability Layer technology for use in the pilot.
Interoperability	Use cases in InterConnect are about interoperability of devices and services, interconnected by the (electrical) grid and the cloud. Technical/communication interoperability is achieved by making use of standardised semantic web technology. Semantic interoperability is achieved by having a shared understanding of (semantic) concepts as defined in the InterConnect set of ontologies. A quintessential part of this set is the Smart Applications REFerence (SAREF) ontology suite as defined by ETSI. It is investigated if adaptations of SAREF are needed for better support of certain use cases. Also, standardized information models from EEBUS and CENELEC are used to create interoperability in energy flexibility services.

2.8 SMARTBear IoT research project

2.8.1 Overall characteristics

Table 19 – SMARTBear characteristics

Reference	https://www.smart-bear.eu/
Description	The primary goal of SMARTBEAR project is to develop an integrated platform gathering numerous health related data flows, to analyse the day-by-day patients' activities and their health status. These analyses are then used to provide evidence-based, personalized interventions towards improving the degree of healthy and independent living of the patients.
Stakeholders and concerns	Consortium partners form a synergy of the health sector and IT technologies: large hospitals owning large patients' datasets, IT companies providing secure collection, storage and analysis of the data, medical devices vendors. Concerns: lack of centralised big data platforms to analyse and process the medical data in a privacy-preserving manner, to prevent the development/deterioration of various conditions of the elderly and reduce medical costs.
Architecture principles	

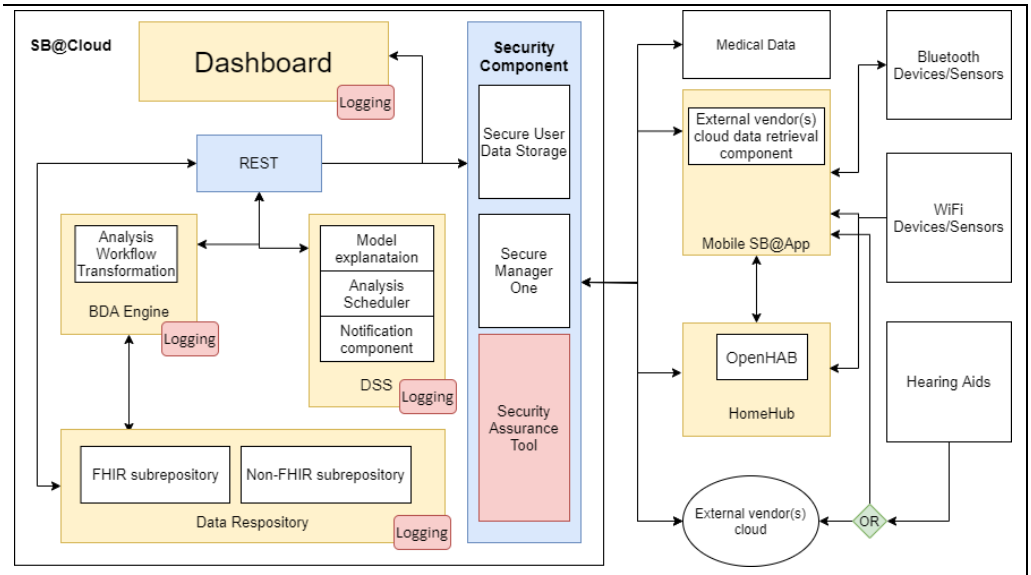


Figure 24 – SMARTBear architecture

The overall system is divided into three main parts – the Mobile SB@App (top-right), the HomeHub (bottom-right), and the SB@Cloud (left) subsystems. The first two are used to co-ordinate the collection of personal data from different devices, and transmit them to the latter subsystem, where the data are anonymised, stored, analysed, and then suggested personalised interventions are produced by the Decision-Support Subsystem (DSS) and transmitted back to the Mobile SB@App, which runs on a smartphone operated by each patient, to inform the patients about actions they could take to improve their health.

Medical data in the SB@Cloud are held in a FHIR-compliant repository. There are also medical data that come from external sources (e.g., hospital EHR systems) and systems developed by other EU research projects (Smart4Health [28], Holobalance [29]).

Data management capabilities

Following features are supported:

- Access control and Authentication: security component consisting of secure assurance tool, secure manager one and secure user data storage
- Federation (partially): through data models that follow the FHIR standard and in one direction only (data digestion – no data exporting).
- Privacy-by-design: all data are pseudo-anonymized before storage by replacing their original person ID with another one that has been allocated to that person from the project upon their registration. In this manner the original data provider (e.g., hospital EHR) cannot link back a person's data to that person.

Roadmap

Full data federation (exporting data to other providers/projects as well) is being considered at the moment. The validation on the large scale pilots will be finished by 2024.

[28] https://smart4health.eu
[29] https://holobalance.eu/

2.8.2 Integration of IoT and Edge Computing

Table 20 – Integration of IoT and Edge in SMARTBear

Connection to IoT and Edge (a.k.a. computing continuum support)	There are studies run across different countries (Portugal, Spain, France, Italy, Romania, and Greece) that will integrate and validate the various tools related to the collection, storage, and analysis of the data. Collection will use IoT devices. However, the analysis will be done in the Smart Bear cloud (SB@Cloud in the architecture) and not at the edge (Mobile SB@App). The Smart-Bear cloud platform gathers the data from IoT devices using the SmartBear mobile application (SB@App) running on the user's smartphone and the HomeHub device (placed in their domicile). All information is transferred to the FHIR-compliant database that is part of the main SB@Cloud subsystem, to be analysed and acted upon. The platform will be tested on the full-scale pilots by 2024.
Data governance across computing continuum	Data governance is used in order to ensure the data is available to all pilots. The data is transferred via REST API functions to ensure the data is sent reliably. Also, the data governance is supported by the SmartBear cloud solution itself.
Trustworthiness across computing continuum	The SmartBear platform processes sensitive data. The reliable data communication channels via REST API services are developed to support the transfer of the high frequency data with certain level of reliability. The first version of the platform has been finished and is tested on the first pilot currently.
Interoperability	The SmartBear platform is compatible with other EU projects collecting medical data (Smart4Health [30] & Holobalance [31]). SmartBear platform stores the data received from the other projects using FHIR-compliant repository for interoperability purpose. At the moment, data exchange is in one direction only (digestion – there's no data exporting).

2.8.3 Use case

Table 21 – SMARTBear use case

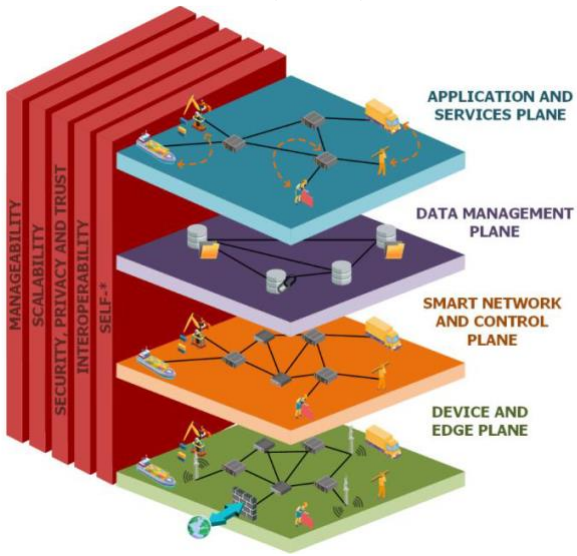
Use case description	The project includes 6 large scale pilots (Portugal, Spain, France, Italy, Romania, and Greece). The SmartBear platform will provide the mechanisms for secure collection, storage, and analysis of the medical data to cover 5 comorbidities: - Hearing Loss - Cardiovascular Diseases - Cognitive Impairments - Mental Health Issues - Balance Disorders - Frailty
IoT and edge systems involved	Collection will use IoT devices. However, the analysis will be done in the Smart Bear cloud (SB@Cloud in the architecture) and not at the edge (Mobile SB@App). The Smart-Bear cloud platform gathers the data from IoT devices using the SmartBear mobile application (SB@App) running on the user's smartphone and the HomeHub device (placed in their domicile). All information is transferred to the FHIR-compliant database that is part of the main SB@Cloud subsystem, to be analysed and acted upon.
Data governance and trustworthiness	The data is transferred via REST API functions to ensure the data is sent reliably. Also, the data governance is supported by the SmartBear cloud solution itself. Security and privacy are supported via security components shown in the middle of the architecture, that is, the Security Assurance Tool, the Secure Manager One, and the Secure User Data Storage.

[30] https://smart4health.eu
[31] https://holobalance.eu

2.9 ASSIST-IoT research project

2.9.1 Overall characteristics

Table 22 – ASSIST-IoT characteristics

Reference	ASSIST-IoT (Architecture for Scalable, Self-*, human-centric, Intelligent, Secure, and Tactile next generation IoT) https://assist-iot.eu/ https://assist-iot.eu/wp-content/uploads/2022/02/ASSIST-IoT_D3.6_Architecture_Definition_Intermediate_v1.0.pdf
Description	ASSIST-IoT aims at design, implementation and validation of an open, innovative multi-plane (semi-) autonomous decentralized reference architecture, associated enablers, services and tools, to assist human-centric applications in multiple verticals. The architecture will support continuous integration and long-term sustainability of domain-agnostic, interoperable, self-* capable, intelligent, distributed, scalable, secure and trustworthy IoT ecosystems.
Stakeholders and concerns	IT Departments of medium/big company (whichever sector), Innovation Departments of a medium/big company or a Research Centre, Head of Development department of a medium/big city. Barriers: Obsolescence of equipment, interoperability among different servers and devices, size and capacities of their diverse computing equipment, incompatibility of solutions.
Architecture principles	The ASSIST-IoT conceptual architecture is rooted in a multidimensional approach, in which Horizontal planes are intersected by Vertical blocks, allowing for a high level of modularity.  **Figure 25 – ASSIST-IoT conceptual architecture**
Data management capabilities	Data Management features supported designed as enablers: - Semantic repository enabler - Semantic translation enabler - Semantic annotation enabler - DLT communication enabler - Long-term data storage enabler - Edge data broker Security, Privacy and Trust enablers: - Identity Manager enabler (OAuth2, Federated identity, W3C VCs) - Authorisation enabler (implementations for XACML) - Security monitoring and threat detection enabler (Wazuh) - Logging and Auditing enabler (IDS (Blockchain-based) Clearing House, Hyperledger Fabric Chaincode (Smart Contracts), cryptographic techniques) - Data integrity Verification enabler, Distributed Broker service enabler (IDS Clearing House, Hyperledger Fabric Chaincode, cryptographic techniques) - DLT-based Federated Learning enabler (Hyperledger Fabric clients - light nodes, openDSU), DAG (IoTa), cryptographic techniques)
Roadmap	The first release of the project is envisioned for April 2022. This (platform-level) release will contain a functional version of the essential enablers, namely: smart orchestrator, long-term storage enabler, edge data broker, VPN enabler, tactile dashboard. The second release is expected for M27 (January 2023).

2.9.2 Integration of IoT and Edge Computing

Table 23 – Integration of IoT and Edge in ASSIST-IoT

Connection to IoT and Edge (a.k.a. computing continuum support)	In the project is 3 main pilots with several subcases are designed, where among different enablers. The functionalities for data management and security, trust and privacy will be implemented and validated. DevOpsSec approach is used for integrating security into the software development lifecycle. This includes automating security testing and integrating security into the Continuous Integration/Continuous Delivery (CI/CD) pipeline to bring security to the process. GitLab is the main code repository for the project with open source licensing. The first release of the project is envisioned for April 2022 for data management selected enablers: long-term storage enabler, edge data broker. The second release is expected for January 2023 with all designed enablers.
Cyber physical systems and digital twins	The first release of the project for security and data management selected enablers supported by DLT and AI algorithms. The second release is planned with other designed functionalities (January 2023). No digital twins approach.
Data governance across computing continuum	The Data Management Plane proposes specific enablers to process, share and present the data, with focus on its semantics.
Trustworthiness across computing continuum	In ASSIST-IoT, privacy and trust per design will be addressed by the introduction of DLT-related enablers.
Interoperability	The implementation of the project is about to take place in three different pilots. Considering this, interoperability will play an especially important role in the fruitful completion of each of the pilots, regardless the systems used in each separate case. Interoperability will be undertaken at three levels: - Technical interoperability – means the ability of two or more information and communication technology applications, to accept data from each other and perform a given task in an appropriate and satisfactory manner without the need for extra operator intervention. - Syntactic interoperability – allows two or more systems to communicate and exchange data in case that the interface and programming languages are different (e.g. by using of a standardisation of the communication between a software client and a server). - Semantic interoperability – is the highest level of interoperability which denotes the ability of different applications/artefacts/systems/… to understand exchanged data in a similar way, implying a precise and unambiguous meaning of the exchanged information. Interoperability will be addressed in terms of scalability, security, privacy and heterogeneity of data sources. ASSIST-IoT will support data interoperability by proposing a semantic data governance toolset, offering data sharing, privacy, security and trust enablers. Another possibility to support the interoperability approaches in ASSIST-IoT is the adoption of DLT.

2.9.3 Use case

Table 24 – ASSIST-IoT use case

Use case description	D3.2: https://assist-iot.eu/wp-content/uploads/2021/12/ASSIST-IoT_D3.2_Use-Cases-Manual-Requirements-and-Business-Analysis-Initial_v1.0.pdf 3 large pilots: - Port automation: ○ Tracking assets in terminal yard (data with time and position of communication / identification) ○ Automated cargo equipment cooperation (Location of the truck, work order information, LIDAR information, Movement recommendations) ○ Remote control with AR support (Video feed from RTG camera system, the location of containers, TOS work orders, commands for moving crane parts (hoist, gantry, straddle). - Smart safety of workers: ○ Occupation safety and health monitoring (Location and proximity data, physiological parameter measurements, weather conditions measurements, personal identification information, training and medical records, building information, users' thermal comfort preferences, alerts and notifications) ○ Fall arrest monitoring (Identity of the user, status of the fall arrest detector, location of the incident) ○ Safe navigation (Location data, navigation instructions along predefined or dynamically updated routes) ○ Health and safety inspection support (Work briefs, safety procedures, required PPE for each type of activity at any location, workers' training records) - Cohesive vehicle monitoring and diagnostics: ○ Fleet in-service emission verification (Sensor measurements, at very high sampling frequencies describing the vehicles' operation and drift correction model parameters) ○ Vehicle diagnostics (Sensor measurements at very high sampling frequencies and thousands of parameters describe the vehicles' operation) ○ Vehicle exterior condition inspection and documentation (High-resolution images, 3D point clouds and corresponding metadata and annotations)
IoT and edge systems involved	Different types of sensors and devices: - Cameras - High frequency sensors - VR devices - Tracking devices
Cyber physical systems and digital twin involved	No digital twins Systems for data management, security, trust and privacy using DLT and AI algorithms.
Data operations	Data from edge devices and sensors as well as other enablers functionalities will be available using Open API, for developers (Open Calls)
Data governance and trustworthiness	Data need to be controlled and protected, especially with privacy protection, trusting requirements for data sources and sharing, processing of data at the edge and in the cloud.

3 Recommendations for standardisation

This report has provided an analysis on the integration of IoT and edge computing in data spaces. Three recommendations are made:

The first recommendation is to agree on data space principles. This paper has identified 12 principles, detailed in Table 1 and summarised in Table 23.

Table 25 – Twelve data space principles

Principles	
1	Data spaces are ecosystems of systems
2	Data usage require provisioning from connecting devices
3	Data spaces support data lifecycle
4	Data interoperability enabled by a common language
5	Data usage enabled by common data models
6	Data curation
7	Trust in data sharing & Data Sovereignty
8	Governance for ethical usage of data
9	Decentralisation
10	Integrated data management
11	Extensible data spaces
12	User-centricity

The second recommendation is to work on data space standards following an architecture of standard as showed in Figure 23:

- Use cases justify concepts, processes, architectures, interoperability, systems. Data space use case standards provide an inventory of application needs that can be used to justify other standards.
- Concepts enable specification of processes, architectures, interoperability, systems. Data space concept standards (terms, principles, ontologies, frameworks) can be used to enable the specification of process, architecture, interoperability and system standards.
- Processes support the construction of architectures, interoperability, systems. Data space process standards can include management, engineering, conformity assessment, continuity management.
- Architectures and Interoperability support the construction and integration of systems
    - **Data space architecture standards** provide means to build high-level specifications of systems.
    - **Data space Interoperability standards** provide means to enable exchange of information, or service provisions by systems.
- System support integration into systems of systems, ecosystems. Data space system standards provide means to construct and operate systems of systems (ecosystems)

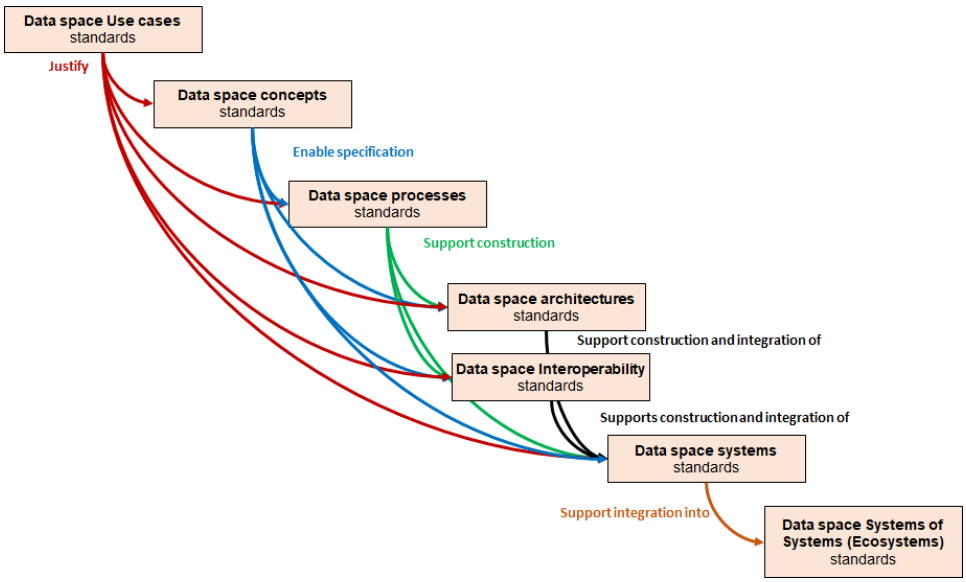


Figure 26 – Potential data space standards

The third recommendation is to integrate IoT, Edge and digital twin concerns in data space standards. Note that the standards should be jointly worked out by working groups focusing on AI, on data, on data governance, on IoT, on CPS and on digital twins.

Contributors

Editor:

Antonio Kung (Trialog)

Reviewers:

Damir Filipovic (AIOTI)

Contributors:

Antonio Kung (Trialog)

Gyu Myoung Lee (Liverpool John Moores University, Professor)

Joakim Koss (JK Consulting)

Georgios Karagiannis (Huawei)

Marco Carugi (Huawei)

Natalie Samovich (Enercoutim)

Philippe Sayegh (Verses Global)

Asbjorn Hovsto (Hafenstrom)

Acknowledgments

This report was prepared by the standardisation working group of AIOTI chaired by Georgios Karagiannis (Chair) and Antonio Kung (co-chair), as part of a joint undertaking with BDVA[32] that will produce another position paper (Data sharing spaces and interoperability)

Contributions concerning IDSA, oneM2M, ETSI MEC as well as the European pilot projects PLATOON, INTERCONNECT, SMARTBear, ASSIST-IoT are acknowledged.

Contributions have also been provided by invited experts: Gyu Myoung Lee and Joakim Koss.

[32] https://www.bdva.eu/

About AIOTI

AIOTI is the multi-stakeholder platform for stimulating IoT Innovation in Europe, bringing together small and large companies, start-ups and scale-ups, academia, policy makers and end-users and representatives of society in an end-to-end approach. We work with partners in a global context. We strive to leverage, share and promote best practices in the IoT ecosystems, be a one-stop point of information on all relevant aspects of IoT Innovation to its members while proactively addressing key issues and roadblocks for economic growth, acceptance and adoption of IoT Innovation in society.

AIOTI's contribution goes beyond technology and addresses horizontal elements across application domains, such as matchmaking and stimulating cooperation in IoT ecosystems, creating joint research roadmaps, driving convergence of standards and interoperability and defining policies. We also put them in practice in vertical application domains with societal and economic relevance.