CREATE-IoT

H2020 Work Programme 2016-2017

H202-IoT-02-2016

Duration: 36 months

European Large-Scale Pilots Programme

CROSS FERTILISATION THROUGH ALIGNMENT, SYNCHRONISATION AND EXCHANGES FOR IoT

H2020 – CREATE-IoT Project

Deliverable 06.06

Final report on IoT standardisation activities

Revision: 1.00
Due date: 30-04-2020
Actual submission date: 04-05-2020
Lead partner: ERCIM

Dissemination level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Summary					
No and name	**D06.06 Final report on IoT standardisation activities**				
Status	Released	**Due**	m40	**Date**	30-04-2020
Author(s)	D. Raggett (ERCIM), E. Darmois (ETSI), O. Vermesan (SINTEF), R. Bahr (SINTEF), M. Serrano (NUIG), A. Kung (TI), P. Annicchino (AS), S. Ziegler (MI)				
Editor	Dave Raggett (ERCIM)				
DoW	Final report on the coordinated IoT standardization activities.				
Comments	The work has been carried out within task T06.02 (Pre-normative and standardisation activities) and is the third out of three deliverables from this task dealing with the standardisation aspects of the LSP Interoperability Framework developed by WP06. The task coordinates the activities with the AIOTI WG on standardisation, SDOs and other various IoT Global Alliances for the validation in usage context of most promising standards and gap analysis identification.				

Document history			
Rev.	**Date**	**Author**	**Description**
0.00	08-02-2020	ERCIM	Initial outline structure
0.01	06-04-2020	ERCIM	Revised structure and questions
0.02	21-04-2020	ETSI	New document organisation and additional content
0.03	21-04-2020	ERCIM	Info on databases etc.
0.04	22-04-2020	ETSI	Additional content on 3D Reference Architecture. Editing.
0.05	26-04-2020	ERCIM	Additional content and editing of section 8
0.06	26-04-2020	ETSI	Restructuring ToC and sections 3 and 4.
0.07	28-04-2020	SINTEF	Contributed on section 6.
0.08	28-04-2020	ERCIM	Additions to section 8 (privacy-based business models)
0.09	29-04-2020	ETSI	Contribution to section 9. Some editing.
0.10	29-04-2020	SINTEF	Additions to section 5, 6, and 7.
0.11	30-04-2020	ETSI	Minor edits
0.12	01-05-2020	ERCIM	Revisions to section 3 and work on the 3D Reference architecture
0.13	02-05-2020	ERCIM	Finalisation of content for all sections except 9
0.14	03-05-2020	ERCIM	Finalisation of content for all sections except 9
0.15	03-05-2020	ETSI	Alignment on D06.03 and associated edits
0.16	03-05-2020	SINTEF	Internal review and comments considered.
1.00	03-05-2020	SINTEF	Final version released.

Table of contents

Figures

Tables

1 EXECUTIVE SUMMARY

1.1 Publishable summary

A major challenge for the Internet of Things (IoT) is to enable a large range of innovative services based upon the connection through the Internet of a large set of applications that use the data produced by a variety of sensors and actuators in very different contexts, domains and business models. To reap the expected benefits of the IoT, enable easy deployment of applications, reducing the costs and risks, and providing the confidence and trust, the IoT ecosystem will require that interoperable platforms and technologies be available to the IoT systems designers and developers. These platforms and technologies will be highly relying on standards.

The purpose of this document is to assess the status of relevant standardisation across the IoT domains represented by the IoT Large Scale Pilots (LSPs), to summarise how the technology and standards landscape has shifted during the course of the LSPs, to review the experience gained and to look at how work done by the LSPs can feed into future standards.

This report starts by a survey of the work done by the LSPs regarding standardisation. In order to position this work with the evolving IoT ecosystem, a summary of the LSP Interoperability Framework, and in particular the 3-dimensional Reference Architecture model, is presented. This 3D model is then used to show what are the standards available or under development, and those for further study in each of its dimensions.

The expected benefit of the proposed approach is to propose a way to help the various stakeholders involved in the development of IoT systems (across the whole IoT system lifecycle) identify the support they can get from standards.

More information on the “LSP Interoperability Framework” can be found in the companion deliverable to this document: CREATE-IoT Deliverable D06.03 “Assessment of convergence and interoperability in LSP platforms”.

1.2 Non-publishable information

None, the document is public.

2 INTRODUCTION

2.1 How to use this document

2.1.1 Scope and purpose

This document surveys the IoT standardisation landscape in the context of the work done in the EU IoT LSPs and the relevant technology trends that have emerged over the last three years.

Its purpose is to assess the status of relevant standardisation across the IoT domains represented by the IoT Large Scale Pilots (LSPs), to summarise how the technology and standards landscape has shifted during the course of the LSPs, to review the experience gained and to look at how work done by the LSPs can feed into future standards.

This report proposes an approach to address the complex landscape of IoT standardisation and how the work of the LSPs fit in and contribute to it. Beyond the presentation of the work done by the LSPs, it relies on the LSP Interoperability Framework, and in particular the 3-dimensional Reference Architecture model developed by the LSPs within the LSP Activity Group 02 ("Standardisation, Architecture and Interoperability"). The 3D model is used to show what are the standards available or under development, and those for further study in each of its dimensions. The last part of the document takes a look at emerging technology trends and their implications for what standards will be needed to facilitate open ecosystems of services building upon the data provided by the IoT.

The expected benefit of the proposed approach is to propose a way to help the various stakeholders involved in the development of IoT systems (across the whole IoT system lifecycle) identify the support they can get from standards.

2.1.2 Target group for this document

The target group for this document is the community of people that have to address the definition of the LSPs from inception to implementation, and in particular regarding the support they can get from the IoT community on pre-normative and standardisation to close the main gaps:

- The identification of pre-normative activities of interest for the LSPs.
- The identification of the main standards on which the LSP implementation has been based.
- The contribution of the LSPs to the resolution of gaps via contributions to IoT standardisation.
- The use of the LSP Reference Architecture model to identify the available standard support.
- The characterisation of emerging standards related to the use of data by IoT systems.

2.2 Contributions of partners

This deliverable is the final deliverable of CREATE-IoT Task 06.02 ("Pre-normative and standardisation activities"). The list below shows the specific contribution of partners to the current deliverable.

ERCIM: As Task Leader and editor of the deliverable, ERCIM has contributed to the definition of the overall content and scope of the deliverable, to the collection and analysis of the information from the LSPs, to the analysis of standards gaps and promising pre-normative activities based on the Activity Group 02 Workshops, and to the review of the deliverable.

ETSI has contributed to the definition of the overall content and scope of the deliverable, to the definition of the IoT Standards Framework, to the synthesis of the support standards based on the results of Activity Group 01 on Use Cases and on the Activity Group 02 Workshops, and to the review of the deliverable.

SINTEF has contributed to the definition of the overall content and scope of the deliverable, provided contributions to various sections of the document based on its involvement in one of the LSPs (e.g. AUTOPILOT), and to the review of the deliverable. SINTEF provided the input to the 3-dimensional IoT Reference Architecture model contributing to the mapping of the IoT platform components to the IoT architectural layers and the definition of crosscutting functions and system properties.

NUIG provided contributions to various sections of the document based on its involvement in one of the LSPs (e.g. ACTIVAGE), and to the review of the deliverable.

TL contributed to the coordination of the pre-normative interoperability activities in LSPs, with an involvement in Activity Group 02. TL also reflected on its participation to activities in the area of active healthy ageing, smart cities and on security and privacy.

MI has provided a contribution to the CREATE-IoT WP06 survey, provided contributions to various sections of the document based on its involvement in one of the LSPs (e.g. SynchroniCity), and to the review of the deliverable. MI continues to contribute to the standardization work of the International Telecommunication Union (ITU), the United Nations agency which also serves as one of the international standards developing organizations (SDOs). MI is involved in Study Group 20 on “IoT and Smart Cities and Communities” and ITU-T Focus Group on Data Processing and Management for Smart Cities and Communities.

AS has contributed to support the interaction with and integration of multi-protocol frameworks and supports the standardization effort related to personal data protection, with a focus on ITU-T, ISO and IEC, and to the review of the deliverable.

2.3 Relations to other activities in the project

The present document has been produced by the CREATE-IoT Work Package 6 " IoT Interoperability and Standardization". WP06 is structured into two complementary tasks:

- Task 06.01 ("IoT Interoperability, standards approaches, validation and gap analysis") focuses on practical topics regarding the implementation of LSP Use Cases.
- Task 06.02 ("Pre-normative and standardisation activities") focuses on the contributions from the LSPs and CREATE-IoT to the IoT standards ecosystem. The present document is a deliverable of this task.

The present deliverable has been developed in Task 06.02 and constitutes its final deliverable. It is one of the two deliverables (D06.03 and D06.06) that present and assess the results of the CREATE-IoT Work Package 6.

The present deliverable is complementary to deliverable D06.03 " Assessment of convergence and interoperability in LSP platforms " (produced in Work Package 6 Task 06.01). Whereas the present deliverable focuses on standardisation aspects, D06.03 focuses on other interoperability aspects, in particular the Interoperability Framework.

The work on standardisation outlined in the current deliverable is making use of the work of CREATE-IoT Work Package 2 ("IoT Large-Scale Pilots Ecosystems Arena for Sharing Common Approaches "), in particular when it comes to Use Cases, open APIs or common methodologies.

This deliverable is addressing some of the issues that are in the scope of WP05 ("IoT Policy Framework - Trusted, Safe and Legal Environment for IoT"). The requirements in terms of security as well as in terms of privacy – whose coverage will ensure trust and user acceptance - are key to the success of LSPs.

3 STANDARDS AND THE INTEROPERABILITY FRAMEWORK

3.1 The LSP Interoperability Framework

This section introduces the importance to realising the potential of the IoT of convergence on a small set of standards for Internet protocols and data formats, and the relationship to the interoperability layers as worked on by ETSI and AIOTI. The 3D Reference Architecture model developed with the help of the LSPs is then introduced as a framework for reviewing the role of standards in each of the LSPs.

3.1.1 The narrow waist for IoT interoperability

Interoperability is key to working systems and requires agreements between elements of a system that may be of very different nature. However, the huge potential for the IoT is being held back by fragmentation into incompatible platforms, standards, and technologies.

The key to rectifying this is to identify where convergence on a small set of standards is necessary and where large diverse set of standards are beneficial. The following figure is taken from the bIoTope project's [21] landscape of IoT standards:

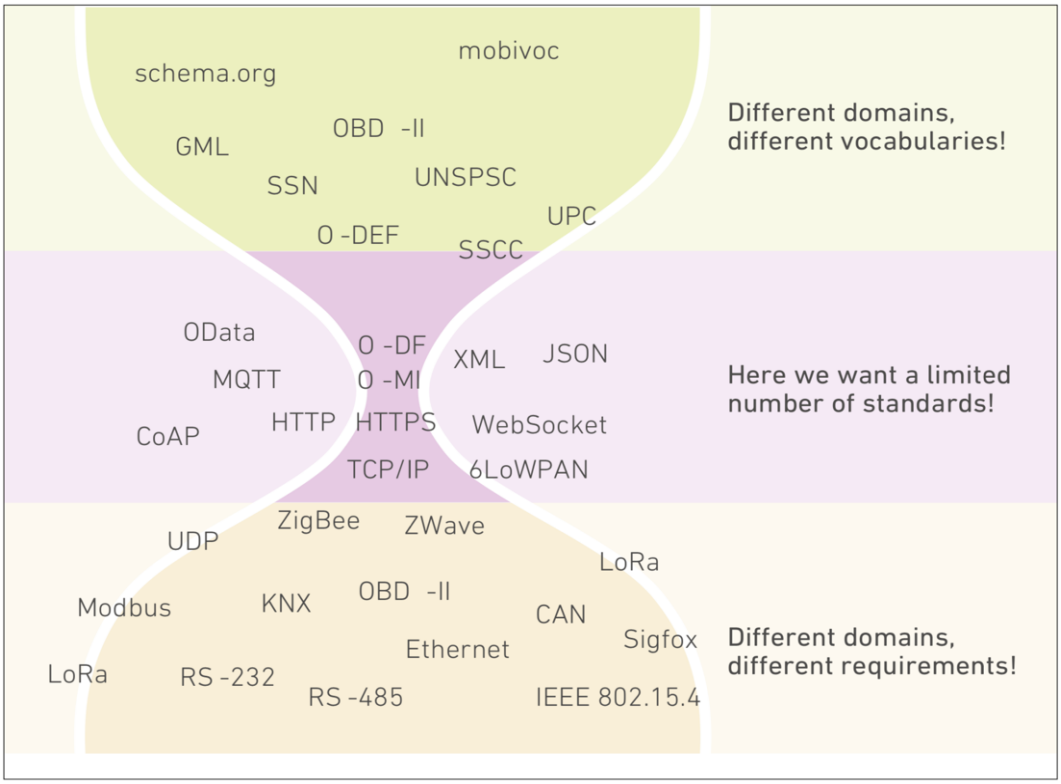

Figure 1: Narrow waist for IoT standards

This illustrates the concept of a narrow waist for the communication protocols that work across IP networks. It is important to note that interoperability relies on more than (for instance) just using HTTP or CoAP. Interoperability cannot be reduced to a means for two systems to exchange information at the network layer. Instead interoperability is best seen through multiple angles: operational behaviour, information exchange, etc. In particular, a layered approach to interoperability has become the accepted paradigm for the description of systems (and this is also true for IoT systems) with a specific focus on four layers:

- *Technical Interoperability* is associated with communication protocols and the infrastructure needed for those protocols to operate.

- *Syntactic Interoperability* is associated with data formats and encodings along with techniques for compressing them.
- *Semantic Interoperability* is associated with shared understanding of the meaning of the exchanged content (information).
- *Organisational Interoperability* is associated with the ability of organisations to effectively communicate and transfer information even across different information systems, infrastructures or geographic regions and cultures.

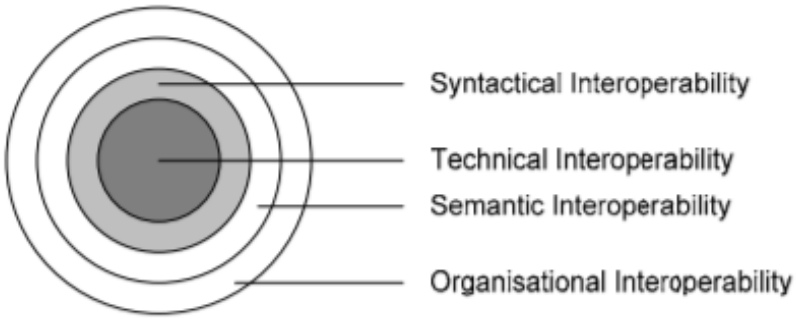

Figure 2: Layers of Interoperability

3.1.2 Main elements of an IoT Interoperability Framework

The IoT community has addressed the question of interoperability in a large number of initiatives during the 2010s. The objective was to rationalise the approach to interoperability by relating it to several elements around which the main design and technology decisions could be taken in the development of an IoT system. The following elements have been gradually emerging as key elements for the rationalisation of

- *Reference Architectures*. In order to achieve interoperability, a lot of elements such as models, definitions or well-defined set of vocabularies need to be agreed upon by the IoT stakeholders in order to ensure a common understanding about the concepts, this is also a preamble to standardisation.
- *Platforms and technologies*. There are hundreds of IoT platforms available for the development of IoT systems. The question of a choice of platform(s) by IoT system designers is complex. Some dimensions have to be considered such as their scope and breadth, the maturity and ownership of their components, and the level of support by standards (and more and more by Open Source).
- *Support of design and development*. For a number of IoT projects, in particular those who span large domains (e.g., Smart Cities), cross-domain interoperability is a key requirement for achieving large scale deployment of IoT-enabled services. On top of a reference architecture model, other elements are required such as cross-application interoperability points (describing where interoperability is supported) and some supporting mechanisms (describing how the support is provided). On top of ad-hoc approaches, project by project, some specifications and standards are emerging to this purpose.

3.1.3 The LSP Interoperability Framework

The first five EU IoT Large-Scale Pilots (ACTIVAGE, AUTOPILOT, IoF2020, MONICA and SynchroniCity) and the associated Coordination Support Action (the CREATE-IoT CSA) have addressed the definition of an LSP Interoperability Framework within the context of their Activity Group 02 on “Standardisation, Architecture and Interoperability”.

The final description of the Interoperability Framework is described in the companion CREATE-IoT Deliverable D06.03 “Assessment of convergence and interoperability in LSP platforms” [3].

The LSP Interoperability Framework has been developed by the Activity Group 02 participants and discussed during a set of Workshops that are summarized in the following CREATE-IoT deliverables:

- "Interoperability Framework Workshop", Deliverable D06.08, 2018.
- "Workshop on LSPs use cases: integration and standardisation alignment", Deliverable D06.09, 2019.
- "Workshop on IoT standardisation activities", Deliverable D06.10, 2019.
- "Workshop on common IoT standardisation framework", Deliverable D06.11, 2020.

During the development of the Interoperability Framework, the following CREATE-IoT deliverables have been produced:

- "Strategy and coordination plan for IoT interoperability and pre-normative and standardisation activities", Deliverable D06.07, 2018.
- "Recommendations for commonalities and interoperability profiles of IoT platforms", Deliverable D06.02, 2018.

The standards and pre-normative activities in support of the Interoperability Framework have also been addressed during its development in the following CREATE-IoT Deliverables:

- "Initial report on IoT standardisation activities", Deliverable D06.05, 2018.
- "Strategy and coordination plan for IoT interoperability and pre-normative and standardisation activities", Deliverable D06.07, 2018.

3.2 A 3-dimensional Reference Architecture Model

3.2.1 A model supporting stakeholders' viewpoints in IoT system development

The LSP three dimensional (3D) Reference Architecture model (developed in the IoT Large Scale Pilots Activity Group 02 "Interoperability and Standardisation") offers an extension of current Reference Architectures and is aiming at ensuring a common view of the different layers of the IoT systems from Physical up to Business; and providing additional viewpoints to the different stakeholders (not just to the developers);

This architecture consists of a 3D representation presenting the key components for IoT/IIoT applications under 3 dimensions that support shared analysis of some between different stakeholders:

- The "Layer" dimension in support of the functional view of the system. The 8 layers defined are common to all the Reference Architectures that the 5 LSPs have developed.
- The "Cross-Cutting Functions" dimension considering transversal technologies such as security, privacy or safety and properties (e.g., integrability)
- The "Properties" dimension.

The applications may require different components presented in the architecture depending on the requirements and specifications.

Two of the 3D Model's dimensions, namely the "Layer" and "Cross-Cutting Functions", provide viewpoints that are present in most of the Reference Architecture model, though in a more systematic approach. The additional third dimension of "Properties" is a new way to discuss the properties of the IoT system between different involved parties (e.g., users, contractors, designers) and identify the elements in support (e.g., functional building blocks, APIs) and those missing.

The 3D architecture is generic and offers a representation that can include the different IoT/IIoT applications across different sector domains (e.g. automated/autonomous vehicles, smart farming, wearables, smart cities, energy, manufacturing, health, etc.). The architecture includes the function by design concept with end-to-end functions addressed across the 8 layers. This allows to address

the heterogeneous applications including different IoT platforms and processing at the edge, fog and cloud.

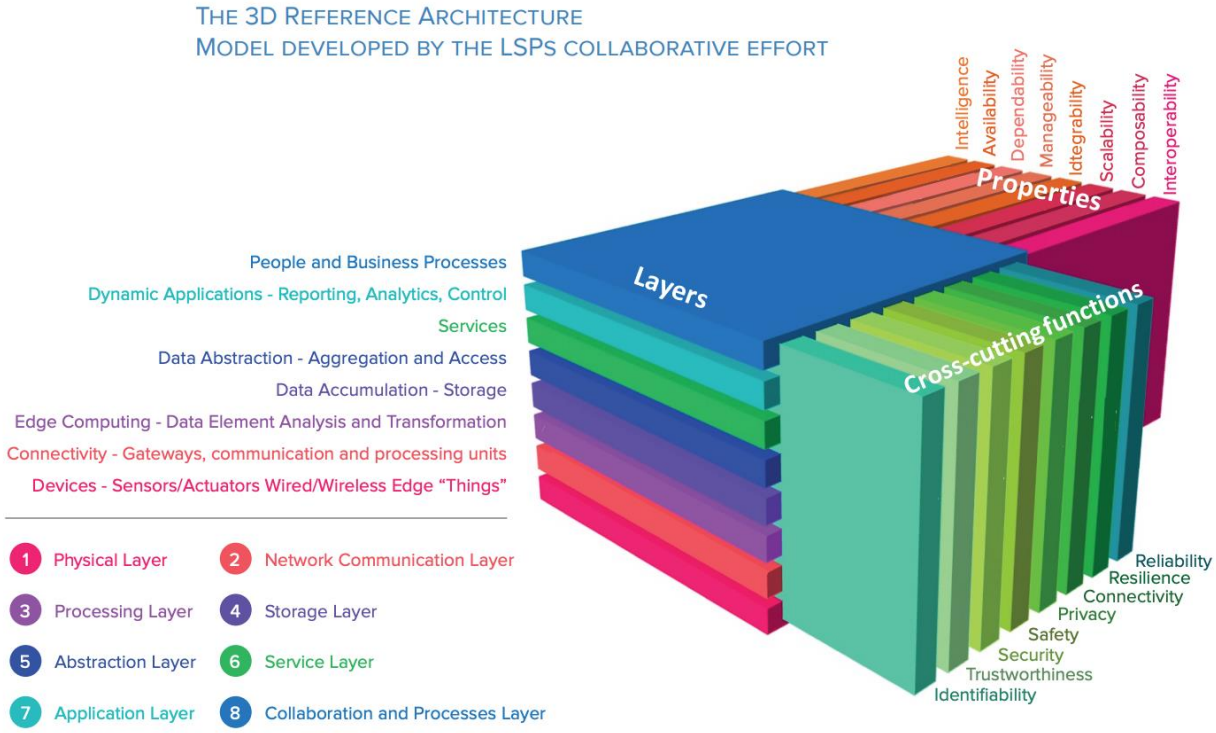

Figure 3: The 3D Reference Architecture Model

This 3D Reference Architecture Model is meant to be fed into the discussion initiated by ISO/ IEC JTC1 on Meta-Architectures.

The 3D Reference Model has been used in the context of the LSP Activity Group 02 to analyse some Use Cases. More of this can be found in CREATE-IoT D06.03 [3].

3.3 Standards and the Interoperability Framework

3.3.1 The importance of Standards in support of Interoperability

Standards are a key element in the IoT Interoperability Framework. A first requirement is to clearly outline the support offered by the current state-of-the-art in standardisation. Beyond this, it is also important to outline the gaps and overlaps (in particular those related to standards): the missing elements of the IoT landscape, mostly due to its complexity, that need to be identified before they may be resolved in the near future.

Pre-normative activities explore promising directions, and just as importantly, attempt to present these in ways that are easy to explain to other communities, thereby helping to build a shared understanding on what new standards are needed.

3.3.2 Standards and the dimensions of the 3D Reference Architecture Model

The survey of standards in the present report is organised in line with the 3 dimensions (Layers; Cross-cutting Functions; Properties) of the 3D Reference Architecture Model developed by the LSPs in conjunction with CREATE-IoT. From a standards standpoint, this model is aiming at:

- Ensuring a common view of the IoT systems from Physical up to Business layers that can allow for the easy identification of the existing standards in support of the functions to be developed at each of the layers of the system architecture.

- Providing additional viewpoints to the different stakeholders (not just for developers) regarding additional cross systems functions such as security, privacy or safety and the identification of associated standards that can be promoted by the specialists of the cross-cutting functions (e.g., security specialists) and incorporated upfront to the design of the system;
- Supporting the shared analysis of some properties (e.g. integrability) between different stakeholders.

As already pointed out, these aspects are addressed in the companion deliverable D06.03 (Initial report on IoT standardisation activities) [3].

4 LARGE-SCALE PILOTS AND STANDARDISATION

This section provides a short introduction to each of the IoT European Large Scale Pilots, the standards and technologies each project have implemented, and their contributions to standards development organisations.

4.1 ACTIVAGE

4.1.1 ACTIVAGE applications and technical setting

The ACTIVAGE project [13] includes 49 partners from across Europe. The project aims to use the IoT to support older people, helping them to stay safe and independent, to have an active social life, and to ameliorate the negative impact of chronic disorders and degeneration. The project built a European IoT ecosystem across 9 deployment sides in 7 European countries (Leeds (UK), Galicia, Madrid and Valencia (Spain), Grenoble, (France), Ober Ramstadt (Germany), Emilia Romagna (Italy), Greece and Finland) with 6000 elderly users and 1200 carers.

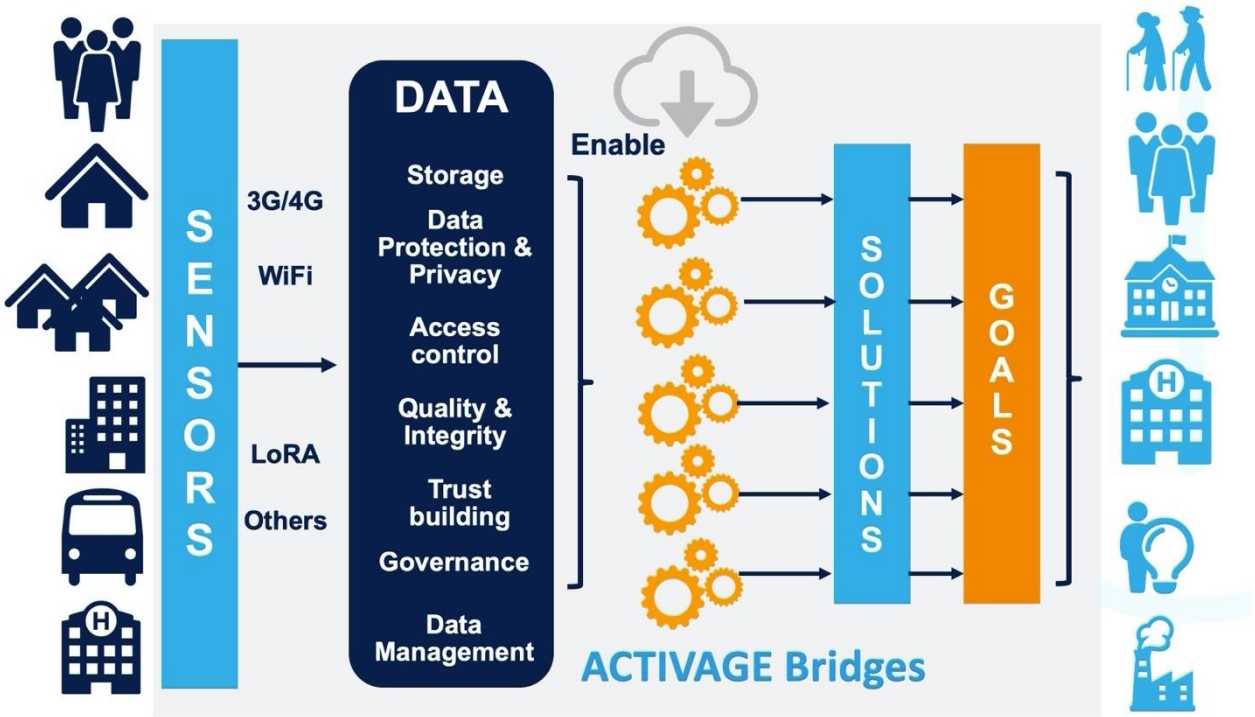

Figure 4: Person-centric IoT Ecosystem (AIoTES)

The project targets people who are classified as managing well, vulnerable, mildly or moderately frail. The aim is to slow the progress of people over time along this axis. The means to support this – and the set of associated applications -are as follows:

- Daily activity monitoring
- Integrated care
- Monitoring assisted persons outside the home
- Emergency trigger
- Exercise promotion
- Cognitive stimulation
- Prevention of social isolation

- Safety and comfort, safety at home
- Support for transportation and mobility

From a technology standpoint, one first goal of ACTIVAGE is to integrate IoT devices (Indoor sensors/actuators, Outdoor sensors, Wearable sensors, Health devices, User interaction devices) with the set of applications.

This is done via the usage of existing IoT platforms (FIWARE, OPENIoT, sensiNact, Sofia2 and universal; see [3] for more details) that implement and embed the ACTIVAGE IoT Ecosystem Suite (AIOTES). These AIOTES include a set of techniques, tools and methodologies that address interoperability, trustworthiness, privacy, data protection and security.

The development of AIOTES in the context of the 7 pilot projects has been a major aspect of the work in ACTIVAGE with an overarching objecting of creating common, replicable solutions across the various implementations. As a result, ACTIVAGE has paid great attention to standards, with an early identification of applicable standards and the development of a number of standardisation activities that have produced a set of contributions to standardisation.

4.1.2 Standards and Technologies Implemented in ACTIVAGE

Table 1 presents the standards that have been used by ACTIVAGE during the deployments on the pilot sites.

Table 1: Standards implemented in ACTIVAGE pilot deployments

Standard	SDO/SSO/Other	Scope
ACTIVAGE taxonomy, vocabulary, DS ontologies	ACTIVAGE	Wellbeing
AHA ontology	ACTIVAGE	Wellbeing
DICOM GSDF	NEMA	Healthcare
EHR R1-2008 (R2014) "Behavioral Health Functional Profile"	ANSI/HL7 Internat.	Healthcare
HL7 (Health Level Seven International)	ANSI/HL7 International.	Security
IEC 60601-1	IEC	Healthcare
IEC 60601-2	IEC	Healthcare
IEC 62304	IEC	Healthcare
IEC 62366	IEC	Healthcare
ISO 13606	ISO	Healthcare
ISO 13606	ISO	Healthcare
ISO 14971	ISO	Healthcare
ISO/IEEE 11073	ISO/IEEE	Healthcare
QUDT (Quantities, Units, Dimensions & Types)	NASA	Sensors
OGC Reference Model	Open Geo. Consortium	Sensors
OpenGIS	Open Geo. Consortium	Sensors
Purl		Sensors
STRIDE		Security
UMLS	NIH	Healthcare
W3C SSN/SOSA ontology	W3C	Sensors

4.1.3 ACTIVAGE Contributions to SDOs

ACTIVAGE monitored and aligned with standardisation groups focusing on active and healthy aging (AHA) and IoT:

- AHA Data model
- ACTIVAGE ontology
- SAREF 4 Health extensions

ACTIVAGE was strongly influenced by work at HL7:

- ANSI/HL7 EHR R1-2008 (R2014) “Behavioral Health Functional Profile”

ACTIVAGE contributed to IEEE P2510 in respect to data quality for IoT sensors, which is enabling a new opportunity for enabling services about certification and validation of sensors quality, following an official certification process defined as part of this standard. At the same time that enables higher reliability and information to AI and data-driven services about the reliability of data sources.

ACTIVAGE supported the cooperation with EC bodies as ENISA in the generation of two whitepapers about IoT security, naming, "Guidelines about IoT Security " and "Guidelines for embedded Software development".

ACTIVAGE also supported data-driven platforms and interoperability with activities around ETSI NGSI-LD in cooperation with FIWARE Foundation, ETSI SAREF and other activities in AIOTI (IERC semantic group) and ITU-T SG20.

4.2 AUTOPILOT

4.2.1 AUTOPILOT applications and technical setting

AUTOPILOT [15] is a three-year innovation action with 45 partners that focused on connected autonomous vehicles. The project developed a range of IoT enabled services which were tested in the main pilot sites (Tampere, Finland; Versailles, France; Livorno, Italy; Daejeon, Korea; Brainport, the Netherlands; Vigo, Spain).

From a technology standpoint, the AUTOPILOT platforms support multiple driving modes:

- *Urban driving* – point to point automated driving in urban environments, requiring vehicles to identify, predict and react in a variety of complex situations.
- *Automated valet parking* – where drivers drop their cars off at a predefined location, and the car is stored and later brought back via the parking management system. Additional optional services include fuelling, recharging, and cleaning.
- *Platooning* – involving a lead vehicle and multiple highly automated or driverless following vehicles, which have automated steering and distance control, using vehicle ton vehicle communication for control.
- *Real-time car sharing* – for commercial and individual sharing services. A fleet management system optimises vehicle allocation in respect to pick-up and drop-down requirements.
- *Highway pilot* – involving processing data from vehicle and road-side sensors to locate and characterise road hazards and provide following vehicles with meaningful warnings and recommendations.

with the following driving services:

- Automated driving route optimisation – with redistribution of traffic along alternative paths towards available parking spots, considering current traffic conditions, the weather, popular events, and available parking spots at the destination.

- *Vulnerable road user sensing* – providing information to autonomous vehicles on pedestrians, cyclists, etc. that are crossing the road ahead, based on data from the devices that such people are wearing, e.g. smartphones, smart watches and smart glasses.
- *Driverless car rebalancing* – where autonomous vehicles move independently or as part of platoons.
- *High definition maps* for automated vehicles – using data collected from connected vehicles.
- *Dynamic eHorizon* – exploiting data from connected vehicles and other sources to dynamically optimise routing.
- *Sixth sense driving* – assessing driving risk metrics based upon a vehicle's on-board sensors, along with specifying metrics and services for road rating.

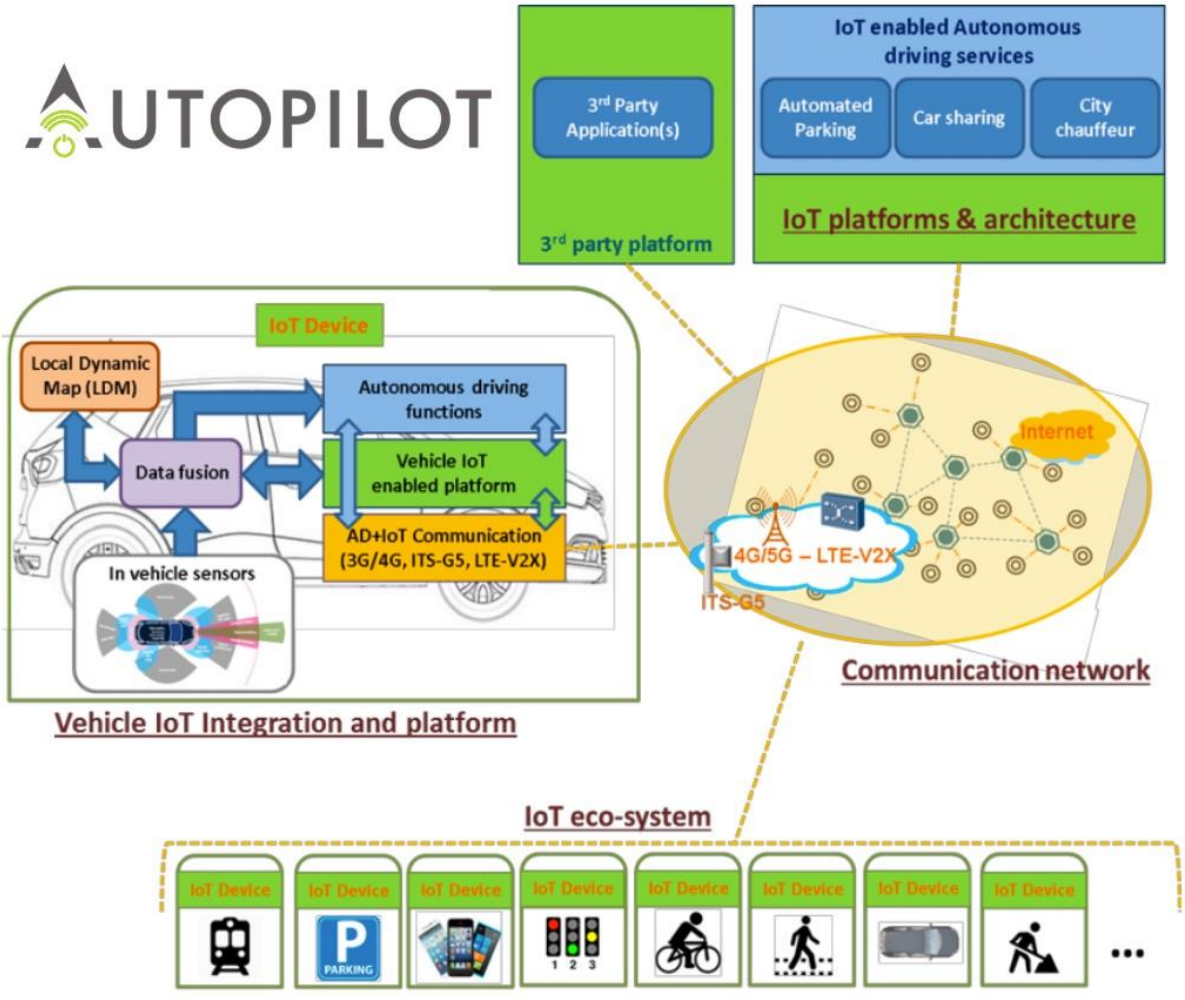

Figure 5: Overall concept for Autopilot

Figure 5 depicts the AUTOPILOT overall concept including the different elements (in particular from an Interoperability Framework perspective) to apply IoT to autonomous driving:

- The overall IoT platforms and architecture, allowing the use of the IoT capabilities for autonomous driving.
- The Vehicle IoT integration and platform to make the vehicle an IoT device, using and contributing to the IoT.
- The Automated Driving relevant sources of information (pedestrians, traffic lights ...) becoming IoT devices and extending the IoT eco-systems to allow enhanced perception of the driving environment on the vehicle.
- The communication network using appropriate and advanced connectivity technology for the vehicle as well as for the other IoT devices.

4.2.2 Standards and Technologies Implemented in AUTOPILOT

The AUTOPILOT standardisation plan addressed:

- The main applicable standards in the IoT and ITS[1] domains as input to architectural and design choices.
- Standardization activities e.g. contributions to standards arising from the development and pilot deployment phases about gaps, inconsistencies, aspects to clarify and new requirements, presentations and / or articles on IoT/AD standards.

The main achievements of AUTOPILOT with respect to standardisation can be found in its Deliverable D5.8 [16].

Table 2: Standards implemented in AUTOPILOT pilot deployments

Standard	SDO/SSO/Other	Scope
3GPP 3G	3GPP	Connectivity
3GPP 4G	3GPP	Connectivity
3GPP 4G LTE-V2X	3GPP	Connectivity
3GPP 4G NB-IoT	3GPP	Connectivity
6LowPAN	IETF	Connectivity
Bluetooth/BLE	Bluetooth SIG, IEEE	Connectivity
CAM (Cooperative Awareness Message)	Open Source, ETSI	ITS
CAN Bus (ISO 11898)	ISO	Devices
CoAP (Constrained Application Protocol)	IETF	Devices
DATEX (Exchange of traffic related data)	CEDR	Data exchange
DDS (Data Distribution Service)	OMG	Data exchange
DENM (Decentralized Environm. Notificat. Message)	Open Source, ETSI	Data exchange
IEEE 802.11-OCB (IPv6 packets over 802.11)	IEEE	Connectivity
IEEE 802.15.4 (Low Rate Wireless PAN)	IEEE	Connectivity
ITS-G5 (ad-hoc V2V communications at 5,9 GHz)	ETSI	ITS
LDM (Local Dynamic Maps, ISO/TS 18750:2015)	ISO	ITS
MQTT	Open Source	Messaging
OGC Reference Model	Open Geo. Consortium	Sensors
oneM2M MCa interface	oneM2M	Service Layer
SPAT / MAP (Signal Phase and Time)	ETSI	ITS

Notes:

- HTTP (Hypertext transfer protocol) is a standard developed by the IETF for the World Wide Web, and commonly used for REST based APIs as well as transfer of Web resources like HTML. HTTPS is essentially HTTP over an encrypted connection between the client and server.
- MQTT is a message broadcast protocol that relays messages via brokers to subscribers for particular message topics. MQTTS uses an encrypted channel.
- 3GPP provides a suite of standards for cellular wireless networks including 3G, 4G and 5G as well as Cellular V2X, which is also known as vehicle-to-everything. This includes vehicle-to-infrastructure vehicle-to-network, vehicle-to-vehicle, vehicle-to-pedestrian, vehicle-to-device and vehicle-to-grid. The main motivation is road safety, traffic efficiency and energy savings.

[1] ITS – Intelligent Transportation Systems

- CAN bus, otherwise known as controller area network, defines a message-based protocol used within vehicles for communication between electronic control units and other devices. It is associated with EOBD/OBD-II on-board vehicle diagnostics standards.
- oneM2M defines a service-oriented framework for transfer of sensor data to client applications over protocols such as MQTT and HTTP, using a three-layer architecture comprising applications, services and networks. MCA defines a REST based interface between applications and services.
- The Internet Engineering Task Force (IETF) is responsible for the core standards for the Internet with IPv4, IPv6 and a suite of protocols including UDP, TCP, HTTP, Web Sockets, CoAP, SMTP, IMAP, RTP and so forth.
- IEEE 802.11 is a suite of standards for Wi-Fi networks. OCB mode provides support for message authentication and privacy via integration of the message authentication code into the operation of a block cypher.
- ETSI ITS G5 is standard for short-range wireless communication designed for intelligent transport systems (ITS) and road transport and traffic telematics (RTTT). CAM is short for "Cooperative Awareness Message" and is used for the exchange of information between road users and roadside infrastructure. DENM is short for "Decentralized Environmental Notification Message" and is used for information related to a road hazard or an abnormal traffic condition. SPaT is short for "Signal Phase and Timing" and is used to convey the status of one or more signalized intersections.
- Bluetooth is a short-range low power wireless protocol that can, for example, be used for wireless headsets for hands free phone calls, and for playing music from a smart phone through a vehicles entertainment system.

4.2.3 AUTOPILOT Contributions to SDOs

During the lifecycle of the project, 25 contributions based on the activity carried out in AUTOPILOT have been submitted to standards development organizations (SDOs).

A significant number of use cases based on AUTOPILOT activity were approved by oneM2M and included in TR-0026 "Vehicular Domain Enablement" [21] and by AIOTI and included in its report "IoT relation and impact on 5G" [23] In more detail:

- Six contributions submitted to oneM2M (five accepted and integrated in TR-0026) and to AIOTI WG3, adding new use cases focused on autonomous driving.
- Participation to the last ETSI ITS CMS Plugtests™ with a vehicular PKI compliant to the new security standards ETSI TS 102 941 v1.3.1 e ETSI TS 103 097 v.1.3.1: compliance and interoperability tests together with 25 stakeholders and 50 observers.
- The PKI by CNIT is available to the project to test secure V2X communication.
- Realization aof the NGSI-LD Context Broker SCORPIO following the ETSI ISG Context Information Management standard. Integration with AUTOPILOT oneM2M platform and interworking with SynchroniCity LSP. SCORPIO will be released as Open Source.

4.3 IoF2020

4.3.1 IoF2020 applications and technical setting

The IoF2020 project [17] has the aim of establishing an IoT ecosystem in the agri-food domain to support smart farming.

The project is unusually large with 120 partners and a duration of four years. Trials have been conducted in five main areas: arable crops, dairy, fruits, vegetables, and meat.

The project seeks to show the value of using high resolution data obtained with ground sensors and drones as compared to remote sensing data obtained from orbiting satellites.

This enables farm machinery to use precise spatial locations together with sensors and actuators, e.g. to measure how yield varies across different parts of a field, and to dynamically tailor delivery of irrigation, fertilizer, pesticides and herbicides to reduce waste, costs and energy expenditure.

Interoperability is a challenge as farm machinery generally uses vendor specific communication.

This was addressed using the Agricultural Data Application Programming Toolkit (ADAPT) and the Extended FMIS Data Interface (EFDI) for a cloud-based API with multi-vendor ADAPT plugins, and support for farmers to sell their data from a data sharing platform.

Livestock sensors can be used to monitor grazing time and location, as well as milk yield from individual cows.

Leg mounted sensors can be used to identify lame animals, as lameness entails pain and discomfort, with decreased fertility and milk yield. Pregnant cows can be given precise tailored doses of mineral supplements.

IoT provides opportunities for improved traceability and enriched information flows along the supply network as well as the means to monitor temperature, humidity and shocks during shipping and storage.

Provenance is of increasing interest to consumers who want to know where and how their food was produced, and to feel a connection to the farms and farmers

The IoF2020 Ecosystem combines sensing and monitoring, analysis, and control, bridging the physical IoT layer to the open IoT architecture and infrastructure.

An online catalogue is being used to show case project results.

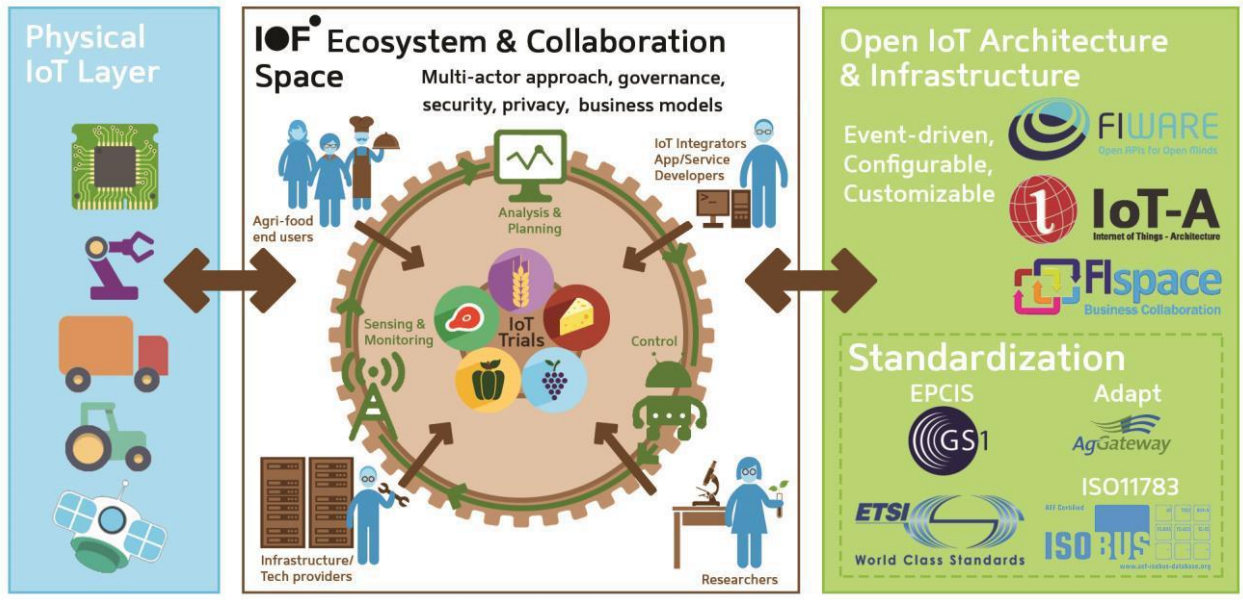

Figure 6: Overall concept for IoF2020

The IoF2020 analysis of gaps and barriers is framed in terms of the layers identified in Figure 7.

This reference model is also used for the identification of applicable standards.

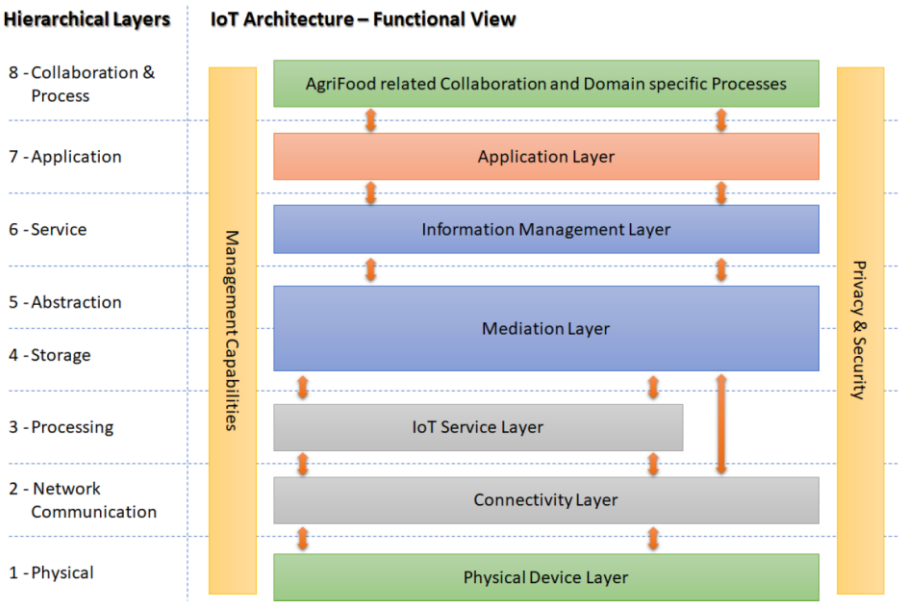

Figure 7: IoF2020 IoT Architecture Reference Model

In addition, the following entities have been highlighted:

- Open Data providers, e.g. data on pests, diseases, historical weather data, and services for contextual data e.g. weather forecasts and flood warnings, and satellite data.
- Harmonised information models to enable interoperability and portability.
- Public GeoServices with geospatial data relating to agricultural assets.

4.3.2 Standards and Technologies Implemented in IoF2020

IoF2020 has defined a number of Interoperability Points (IOPs) which can be exploited to integrate available systems to other systems. These IOPs are related to standard has described in Figure 8.

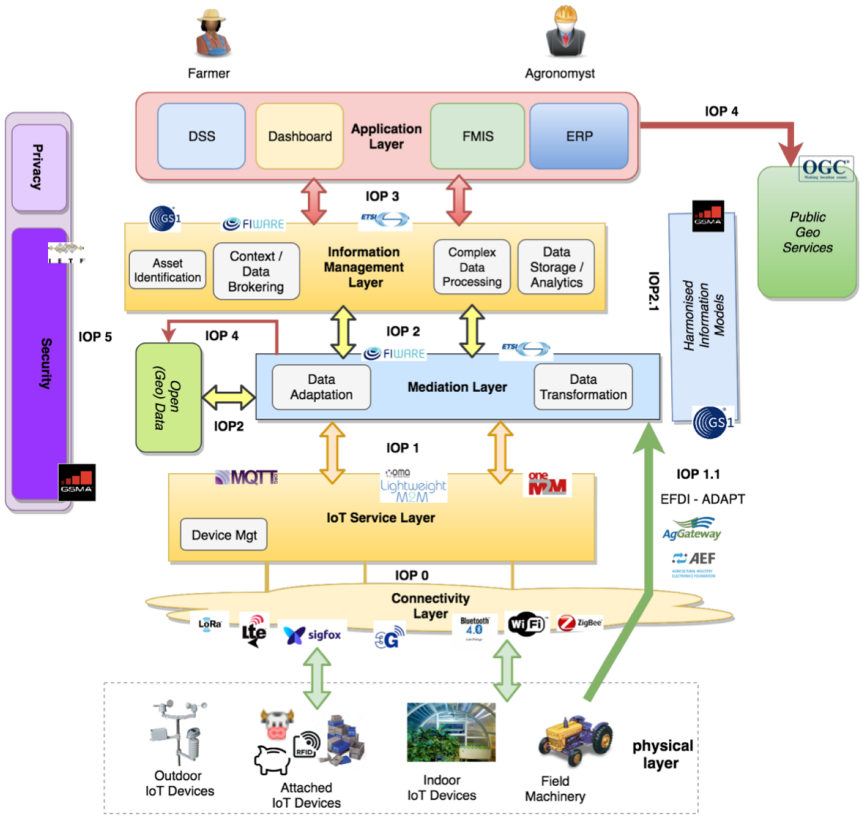

Figure 8: Overview of IoF2020 End Points and their relationship to standards

During the implementation of the IoF2020 use cases, a number of standards, many specific to the Agriculture and Food domain have been used, with the most important summarized in Table 3.

Table 3: Standards implemented in IoF2020 pilot deployments

Standard	SDO/SSO/Other	Scope
365 FarmNet		Farm Management
3GPP NB-IoT	3GPP	Connectivity
3GPP LTE-M	3GPP	Connectivity
6LowPAN	IETF	Connectivity
ADAPT (AG Data Application Programming Toolkit)	Open Source	Data Models
Bluetooth LE	Bluetooth SIG, IEEE	Connectivity
CoAP	IETF	Devices
EFDI (Extended Farm Managmt. Information Syst.)	AEF	Agri. Machinery
FMIS (Farm management Information System)	AEF	Agri. Machinery
GS1 GLN (Global location numbers)	GS1	Data Models
GS1 GTIN (Global trade item numbers)	GS1	Data Models
GS1 GRAI (Global returnable asset identifiers)	GS1	Data Models
GS1 GPC (Global product classification)	GS1	Data Models
GS1 CBV (Core business vocabularies)	GS1	Data Models
GS1 EPCIS (Electronic product codes for capt. & shar.)	GS1	Data Models
IEEE 802.15.4 Low Rate Wireless PAN	IEEE	Connectivity
ISOBUS (ISO 11 783)	ISO/AEF	Agri. Machinery
JSON (JavaScript Object Notation)	IETF	Data Exchange
LLRP (Low Level Reader Protocol)	GS1	Devices
LoRaWAN	LoRa Alliance	Connectivity
LWM2M (OMA Lightweight M2M)	OMA	Device Mgt
MODBUS	MODBUS Organisation	Devices
MQTT (-SN)	Open Source	Messaging
NFC (Near-Field Communication)	NFC Forum	Connectivity
NGSI (Next Generation Service Interface)	OMA, FIWARE	Data models
NGSI-LD	ETSI	Data models
OAuth v2	IETF	Security
OGC Web Feature Service (WFS)	Open Geo. Consortium	Sensors
OGC Web Map Service (WMS)	Open Geo. Consortium	Sensors
QR Code (ISO/IEC 18004)	ISO/IEC	Devices
REST (Representational State Transfer) [29]	Uni. California, Irvine	Web services
RFID (Radio Frequency IDentification)	ISO	Devices
SigFox	SigFox Inc.	Connectivity
SOAP (Simple Object Access Protocol)	W3C	Messaging
TLS (Transport Layer Security)	IETF	Security
Wi-Fi (IEEE 802.11)	IEEE	Connectivity
XMPP (Extensible Messaging & Presence Protocol)	XMPP-IoT	Messaging
ZigBee	ZigBee Alliance	Connectivity

Regarding the data models and vocabularies:

- GSMA IoT Big Data harmonised data model, including entities relevant to smart agriculture.

- AgGateway ADAPT which provides a common reference data model for documenting precision farming operations, open source plug-ins and a framework for developing proprietary plug-ins.

The following points can be noted:

- Physical device layer – the need for standardised low-cost components.
- Connectivity layer – short range communications, e.g. Wi-Fi, Bluetooth and ZigBee (IEEE 802.15.4); long range low power communications (LPWA), e.g. Sigfox, NB-IoT, LTE-M, LoRa and ZigBee (> 1 km); note that cellular networks are less interesting as they often suffer from poor coverage and speed in sparsely populated rural areas.
  - Requirements for new machine to machine standards for agricultural applications, e.g. low bandwidth for environmental data and high bandwidth for drone surveys.
  - Interoperability challenges across vendors, see CREATEIoT D6.03 for details.
- IoT Services layer – where MQTT, CoAP (OMA LWM2M) and oneM2M are the most relevant. The competing standards create burdens due to different message exchange paradigms, protocols, and data encodings. Convergence is needed on a common protocol, message formats, data types and units of measure.
- Agricultural machinery – communication using standards from AEF and AgGateway. A need for interoperable implementations of complex standards (ISOBUS) and further work with ISO and ETSI.
- Information management and mediation layers – these are suffering from a proliferation of proprietary APIs, vocabularies, and incompatible data models. Opportunities for improved standards starting from ADAPT, GDS CBV, GSMA IoT Big Data, and NGSI.
- Application layer – need for improvements in usability and interoperability, and further work on security and privacy, competing interests in regard to data ownership (who owns the data and who benefits from the data).

In conclusion, IoF2020 has shown the importance of reducing fragmentation through convergence on common standards for protocols, data formats and models at all layers. A further challenge is to clarify the business models in respect to data ownership as a basis for open ecosystems and lowering barriers to competition.

4.3.3 IoF2020 Contributions to SDOs

The major contribution has been to GS1 via its role as an IoF2020 partner. The IoF2020 pilots provide important use cases for SDOs, along with the lessons learned on interoperability challenges at all layers (see below).

4.4 MONICA

4.4.1 MONICA applications and technical setting

The MONICA project provided a large-scale demonstration of new and existing IoT applications in the context of cultural performances (e.g. concerts) in open air settings and the associated challenges in terms of crowd safety, security, and sound management.

Open air concerts for popular music generally involves electronic amplification and large speakers to provide adequate sound levels for all of the audience. Unfortunately, the sound can spill over into neighbouring areas, where it is perceived as noise pollution. MONICA was able to demonstrate a 15dB reduction in sound levels outside of the event area through a combination of

passive absorbers and secondary speakers for active sound cancelation, functionally analogous to a larger scale version of noise cancelling headphones.

The technology used to realise this involved a wireless transmission system with very low latency and time jitter.

For management of crowds, CCTV cameras and two kinds of wristbands were used, one for visitors and the other for the event staff.

This allow the MONICA platform to monitor crowd density and flow at different locations for each event. Mobile apps and large event screens are used to redirect visitors to safer areas.

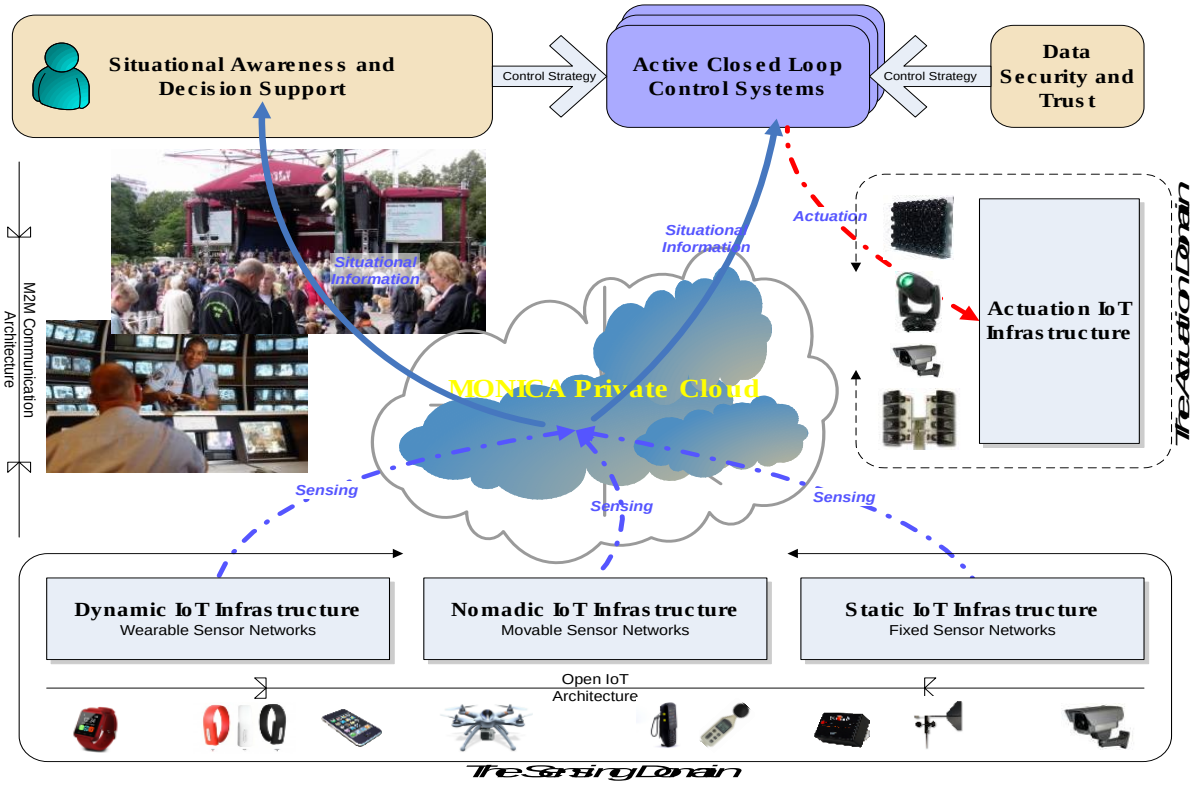

Figure 9: The MONICA overall architecture

Location tracking is supported for staff members using a combination of CCTV cameras, wristbands, and mobile phones. This enables nearby staff to be alerted when incidents need assessment or assistance.

This can also be used to help visitors find friends, or parents to find children who have wandered off. Staff wristbands offer higher precision location than for crowd wristbands and mobile phones.

MONICA deployed pilots in cities across Europe: the “Rhine in Flames” festival and the Pützchens Markt in Bonn, Fredagsrock (Friday Rock) at Tivoli Gardens in Copenhagen, the Port Anniversary and Dom Festival in Hamburg, Rugby matches and cricket matches in Leeds, the Fête des Lumières and Nuits Sonores in Lyons, and the Kappa FuturFestival in Torino.

One challenge was the hurdle of testing novel devices that have yet to be CE-marked. This was addressed through application of Article 9 of DIRECTIVE 2014/53/EU.

This requires devices that are not yet CE-marked to be cleared marked as such.

In addition, the pilot area for the tests should preferably be gated, so that devices are not allowed outside of the pilot area and are to be collected after the event.

A lack of awareness of this directive hindered pilots in some of the countries used by MONICA.

4.4.2 Standards and Technologies Implemented in MONICA

MONICA made use of the standards listed in Table 4.

Table 4: Standards implemented in MONICA pilot deployments

Standard	SDO/SSO/Other	Scope
Bluetooth LE	Bluetooth SIG, IEEE	Connectivity
EN 300 220-2 V3.1.1 (2017-02) for the Crowd Wristbands operating in the frequency range 865 – 868 MHz	ETSI	Connectivity
EN 302 065-2 V2.1.1 (2016-11) for the Staff Wristbands using Ultra Wide Band technology (UWB) in the band 3.4 – 3.8 GHz	ETSI	Connectivity
IEEE 802.11 WiFi	IEEE	Connectivity
IEEE 802.15.4 Low Rate Wireless PAN	IEEE	Connectivity
IETF 6LoWPAN	IETF	Connectivity
IETF CoAP	IETF	Devices
ISO/IEC/IEEE 42010:2011	ISO/IEC	Architecture
LoRa	LoRa Alliance	Connectivity
MQTT	OASIS	Messaging
NFC (Near-Field Communication)	NFC Forum	Connectivity
OGC SensorThings API	Open Geo. Consortium	Sensors
oneM2M	oneM2M	Service Layer
REST (Representational State Transfer)	W3C	Web services
RFID (Radio Frequency Identification)	ISO	Devices
SAREF	ETSI	Semantic Int.
UWB	IEEE	Connectivity
Wi-Fi (IEEE 802.11)	IEEE	Connectivity

More information about how MONICA made use of the above listed standards:

- Bluetooth/BLE – for the wireless connection between a joystick and smart glasses worn by security staff.
- Wi-Fi – IEEE 802.11 – mainly used for sound level meters to send data to their gateway.
- UWB – IEEE 802.15.4a – this standard, which defines both PHY and MAC layers of UWB, has been used as radio interface for the staff wristbands in order to manage security in large open-air cultural events.
- ETSI EN 302 065-2 V2.1.1 for staff wristbands (UWB) - a harmonized standard used for the staff wristbands (based on UWB) mainly focusing on requirements for UWB location tracking.
- ETSI EN 300 220-2 V3.1.1 for crowd wristbands (868 MHz) – harmonized standard that MONICA has used for the crowd wristbands.
- LoRa – used by GPS-based trackers for the security staff to send their positions to a gateway.
- IETF 6LoWPAN / IETF ROLL / IETF CoAP – used in the first version of environmental sensors to send their data to a gateway.
- OASIS MQTT – used by the IoT platform to dispatch messages between the SCRAL module, which is the IoT adaptation layer, and modules running at the Service Layer of the MONICA platform.
- oneM2M – used to integrate in the MONICA platform some environmental data from a Smart City platform made available by the city of Hamburg.
- OGC SensorThings APIs – used for data modelling for the MONICA IoT platform. MONICA has used the GOST server that is a GO implementation of the OGC SensorThings APIs.
- ISO/IEC/IEEE 42010:2011 – to describe the architecture in terms of architectural viewpoints.

4.4.3 MONICA Contributions to SDOs

MONICA's work on sound reduction outside of the event area suggests opportunities for a for a new spectrum optimized IoT radio standard covering "Streaming Optimized" RF links with flow control etc.

This could off-load connections to sensors of all kinds delivering time sensitive streaming information.

This is to be studied further in liaison with ETSI. It is also important to report, that if non-compliant devices are demonstrated, then Art. 9.2 of the RED 2014/53/EU requires that such devices are not brought outside the Pilot area.

4.5 SynchroniCity

4.5.1 SynchroniCity applications and technical setting

The SynchroniCity project [20] aimed to create a global IoT market for cities and businesses to develop shared digital services for smart cities in collaboration with OASC (Open & Agile Smart Cities), itself a non-profit international smart city network connecting over 140 cities in national networks from 27 countries and regions.

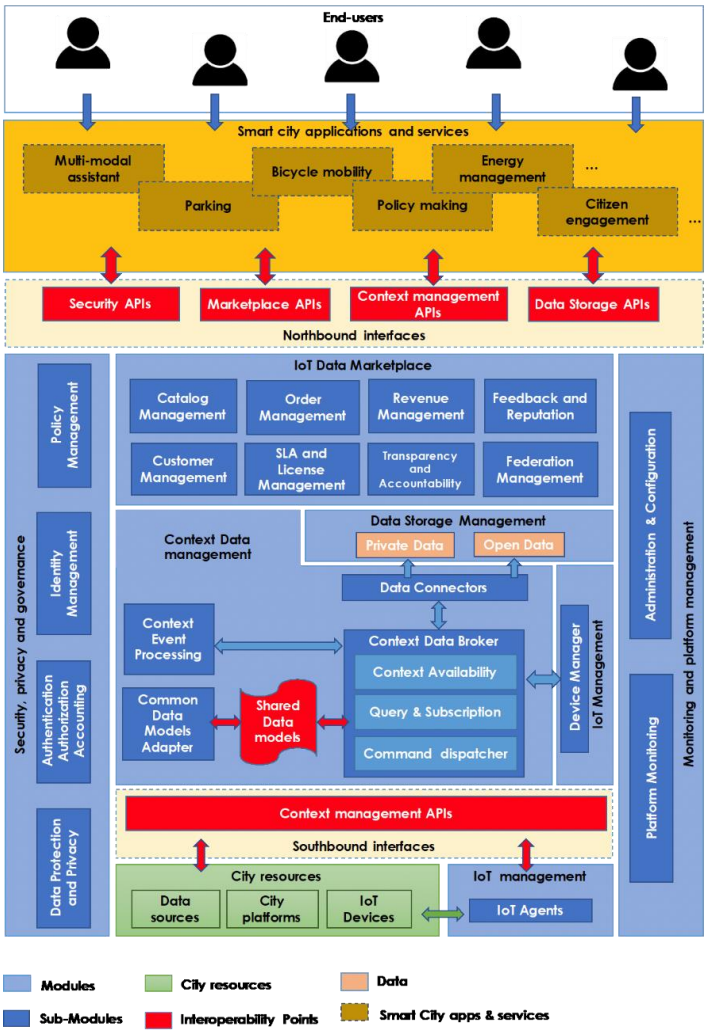

Figure 10: SynchroniCity Architecture and Interoperability Points

SynchroniCity first worked on principles and guidelines that were expressed as an architectural framework model, which itself is a realisation of the OASC "minimal interoperability mechanisms", see Figure 10 where the interoperability points are marked in red.

This starts with the end users who are supported by a variety of applications that are clients of the SynchroniCity platform.

This platform is arranged around a storage system for a marketplace of private and public data and metadata, that in turn is fed by IoT devices and other information sources.

This approach was validated with 49 pilot deployments in 18 cities in Europe and beyond.

The minimal interoperability mechanisms implemented by SynchroniCity include:

- Context information management API with support for detecting specific events and managing the corresponding actions.
- Shared data models across different verticals to facilitate interoperability for applications and systems among different cities.
- Marketplace API with support for catalogues, orders, revenue management, service level agreements and license management.
- Security API for registering and authenticating users and applications.
- Data storage API with access to live and historical data.

4.5.2 Standards and Technologies Implemented in SynchroniCity

SynchroniCity made use of the standards listed in Table 5.

Table 5: Standards implemented in SynchroniCity pilot deployments

Standard	SDO/SSO/Other	Scope
FIESTA (FIESTA-IoT Semantics Library)	EU FIESTA	Semantic Int.
GTFS (General Transit Feed Information)	-	Transportation
Hypercat (hypermedia catalogue format)	Hypercat	Semantic
MQTT	Open Source	Messaging
NGSI	ETSI	Data Model
OASC (Principles and data model)	OASC	Data Model
OAuth v2	IETF	Security
ODF (Open Document Format)	ISO	Data Model

This section focuses on the standards used by SynchroniCity for context management and security. Additional information on the SynchroniCity marketplace and data storage can be found in the SynchroniCity report D2.10 "Reference Architecture for IoT Enabled Smart Cities".

Context Management

SynchroniCity has adopted work on context information by OMA, ETSI and the FIWARE Foundation, involving a framework for describing the physical entities along with the use of HTTP for queries, updates and subscriptions for notifications of updates.

In 2012, The Open Mobile Alliance (OMA) published the Next Generation Service Interfaces architecture for data configuration and management, including context management (NGSI-9, NGSI-10).

The context management component specifies a RESTful protocol for exchanging XML messages over HTTP. This relies on the following meta-model:

- Entities: identifier, type, and a set of attributes.
- Attributes: name, type, value, and optional metadata.
- Metadata: name, type, and value.

The following figure provides an example involving a house and a car.

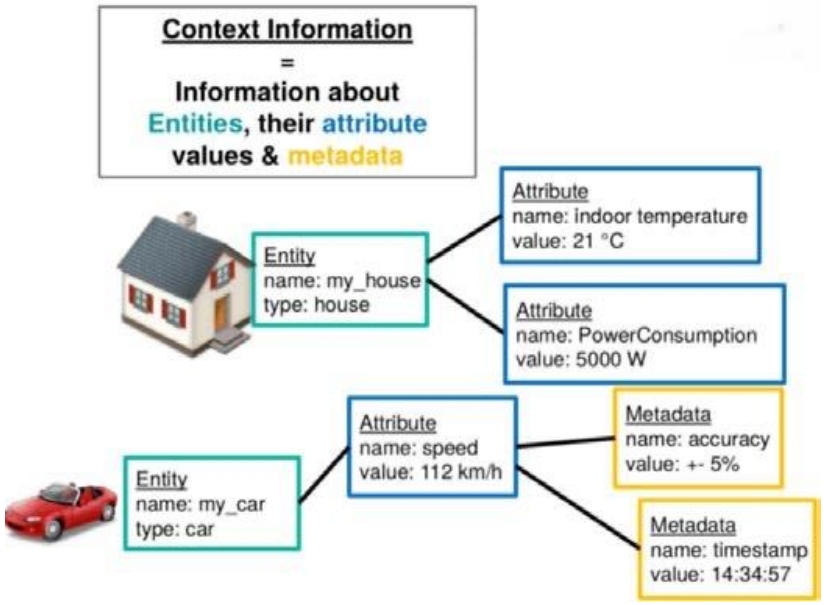

Figure 11: OMA NGSI Context Information in SynchroniCity

FIWARE is an open source IoT platform. The FIWARE Orion Context Manager v1 provides an implementation of the OMA NGSI Context Manager.

The European Telecommunications Standards Institute (ETSI) in close cooperation with the European Commission has pioneered the use of formal ontologies in relation to machine to machine communication, starting in 2015 with the SAREF (Smart Appliances REFerence) ontology for smart appliances to exchange energy related information with an energy management system.

Inspired by OMA NGSI, the ETSI Industry Specification Group (ISG) developed the context information management NGSI-LD API. This uses JSON in place of XML for messages, and assumes a richer meta-modal based upon RDF and Linked Data. It has been implemented as version 2 of the FIWARE Orion context broker.

NGSI-LD distinguishes between context brokers and context registries, where the registries provide information about what context information is available from the brokers. The architecture further allows for federation, e.g. providing a single point of access in the case where different city departments operate their own brokers and registries.

The NGSI-LD core meta-model allows for both entities and their relationships to have properties. This is defined in terms of the W3C definitions for RDF resources, properties, and literals. The cross-domain ontology relates the core meta-model to concepts for measurements, including time and location, drawing upon work by W3C and the Open Geospatial Consortium. This can then be combined with domain specific ontologies, e.g. for an application involving allocation of city parking spaces to cars.

Security

The SynchroniCity security architecture is built upon standards from OASIS (XACML), oneM2M, GSMA and ENISA. Access rights for users are expressed in terms of policies combining attributes such as users, resources, actions, objects, etc. Other access control models can also be supported e.g. role-based access control, in which permissions are based on the roles you embody rather than who you are. Attribute based policies are more general in the sense that they can be expressed in terms of users' identities, users' roles, the time of day, resources, actions to be performed, and so forth.

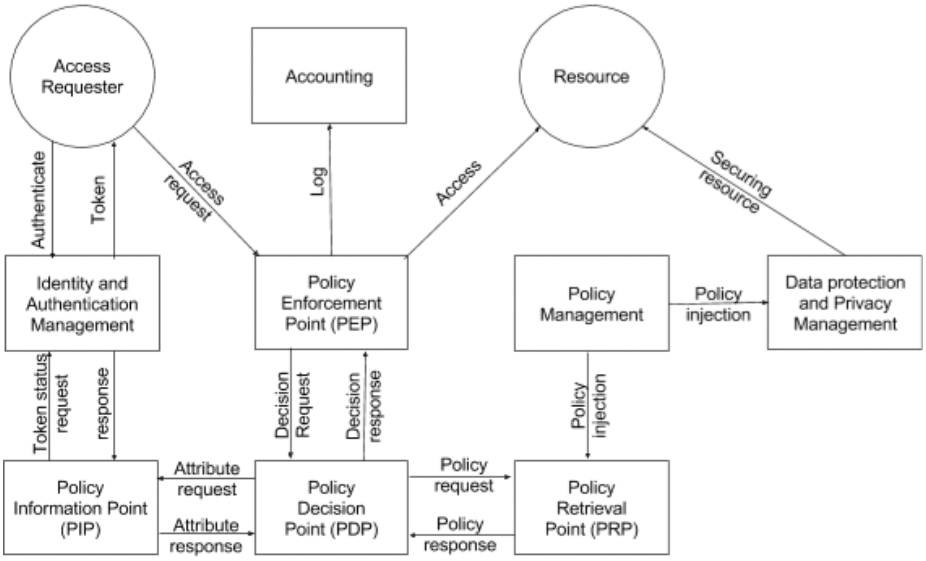

Figure 12: Security Components in SynchroniCity

OASIS XACML (eXtensible Access Control Markup Language) is an XML based access control policy language, architecture and processing model that describes how to evaluate access requests. It can be contrasted with other approaches, e.g. SAML, SPML, OAuth, JWT, OpenID Connect and SCIM.

- *SAML* (Secure Assertion Markup Language) is an OASIS standard for exchanging authentication and authorization information, e.g. between an identity provider and a service provider.
- *SPML* (Service Provisioning Markup Language) is an OASIS standard for service provisioning requests, e.g. requests to add, modify, or delete user accounts, enable or disable access, grant or revoke access rights, change passwords.
- *OAuth* is an Internet standard for granting websites and applications time-limited access to users' information without giving them the users' passwords.
- *JWT* (JSON Web Token) is an Internet standard for creating JSON based tokens that assert one or more claims, e.g. authorisation to use a resource. An example is where a client is authenticated using HTTP and provided with a time-limited token that can be used with messages over Web Sockets to access given services.
- *OpenID Connect* is layered on top of OAuth 2.0 to request and receive information about identities and currently authenticated sessions. OpenID combines the OAuth 2.0 access token with JWT for ID tokens.
- *SCIM* (System for Cross-domain Identity Management) is an Internet standard for managing user identities in cloud-based applications and services.

4.5.3 SynchroniCity Contributions to SDOs

SynchroniCity has supported the convergence of several streams of activities through participation in SDO activities, including ETSI ISG CIM, SF-SSCC, ITU-T FG-DPM, and the organization of several workshops and the coordination of the Smart City track at IoT Week 2017-2019. Furthermore, SynchroniCity is part of an ongoing effort to establish a practical and converging instantiation of the OASC MIMs through the OASC Technology Council.

SynchroniCity has contributed to AIOTI Working Group 3's work on IoT LSP Standard Framework Concepts and to the Dutch standardisation development organisation NEN's "Eerste concept Nederlandse Praktijkrichtlijn voor Open Urban Platforms". SynchroniCity adopted the best practices identified by the European Innovation Partnership for Smart Cities & Communities (EIP-SCC), and has contributed to convergence of smart city standards, through discussions around ESPRESSO H2020 and other smart city projects.

5 STANDARDS AND THE 3D REFERENCE ARCHITECTURE

5.1 Standards for the Layers Dimension

Table 6 is defines and maps the IoT platform components to the eight IoT architectural layers described in this section [11].

Table 6: Mapping of the IoT platform components to the IoT architectural layers [11]

IoT Architectural Layer	Components	Definitions
Physical	Operating system	Offers low-level system SW managing HW, SW and runs applications.
	Modules and drivers	Offers adaptable modules, drivers, source libraries that reduce development & testing time.
	MPU / MCU	Offers multi-purpose programmable electronic devices at microprocessor/microcontroller level.
Network communication	Connectivity Network/ Modules	Offers connectivity networks/HW modules enabling air interface connectivity.
	Edge gateway (HW based)	Offers IoT gateway devices to bridge connectivity from IoT nodes into the cloud-based platform.
Processing	Device management	Enables remote maintenance, interaction and management capabilities of devices at the edge.
	Edge analytics	Capabilities to perform processing of IoT data at devices at edge as opposed to cloud.
Storage	Storage/Database	Cloud based storage and database capabilities (not including on premise solutions).
Abstraction	Event and action management	Simple rules engine to allow mapping of low-level sensor events to high level events and actions.
	Basic analytics	Provides basic data normalization, reformatting, cleansing and simple statistics.
Service	Service orchestration	Supports mashup of different data streams, analytics and service components.
	Advanced analytics	Allows insights from data to be extracted and more complex data processing to be performed.
Application	Visualization	Presents device data in rich visuals and/or interactive dashboards.
	Development environment	Provide integrated development environment to simplify development of apps.
Collaboration/ Process	Business system integration	Enables integration with existing enterprise and other external systems.

5.1.1 Physical Layer

The physical layer (Layer 1 in the OSI model) provides for the transfer of raw bits of information. In virtually all cases, this is achieved electromagnetically via an optical, wired or wireless radio medium.

Fibre optic connections are typically used for long range communication, e.g. the Internet backbone. Wired connections include coaxial cables and twisted pair cables such as the Category 6 RJ45 Ethernet cable used for local area networks in offices. Twisted pair cable is also widely used for KNX building automation networks.

Wireless connections can be categorised by their range and effective bandwidth:

- Very short range, e.g. personal area networks (PAN) such as IEEE 802.15.6.
- Short range, e.g. Bluetooth, ZigBee and IEEE 802.15.4 and 802.15.5.
- Intermediate range, e.g. IEEE 802.11 Wi-Fi networks.
- Long range, e.g. 3G/4G/5G cellular networks and low power wide area networks (LPWAN), e.g. SigFox, LoRa, Weightless and Wize.

Low power connections are likely to have a low effective data rate, except for very short-range connections. This makes them suitable for sensors that need a long battery life or which operate on ambient power, e.g. a small solar electric panel, and which only need to report readings occasionally, e.g. temperature and humidity for environmental sensors.

5.1.2 Network Communication Layer

This corresponds to Layers 2 and 3 in the Open Systems Interconnection Reference Model (ISO 7498, ITU-T X.200). Layer 2 defines the mechanisms used to transport network packets together with identifiers for devices. An example is Ethernet where devices are assigned a 48-bit MAC (media access control) address. Layer 3 is generally used for the Internet Protocol, either IPv4 with 32-bit addresses, or IPv6 with 128-bit addresses. The Address Resolution Protocol (ARP) defined by RFC 826 provides a means to map IPv4 addresses to Layer 3 device addresses. For IPv6 networks, the corresponding protocol is the Neighbour Discovery Protocol (see RFC 4861).

Loosely speaking, Layer 1 and 2 standards are typically defined by the IEEE, whilst Layer 3 standards are defined by the IETF. The Internet is essentially an abstract network that is built on top of, and connects, many Layer 2 networks. Internet protocols are either datagram (UDP) or connection based (TCP). Some Internet protocols of interest to the IoT are as follows:

- Transport Layer Security (TLS) a means to provide an encrypted connection over an unencrypted TCP connection. DTLS is the equivalent for UDP based protocols.
- HTTP (hypertext transfer protocol) – this was originally developed for Web pages but has now become widely used for application programming interfaces based upon request/response pairs which both start with a set of name/value headers. The request starts with a named operation (see below) and a path string it applies to. The response starts with a code describing the success or failure of the operation. A common design pattern is "representational state transfer" or REST with operations on resources such as GET, PUT, POST and DELETE. The name derives from the idea of transferring the complete state with each request and response message. HTTPS is HTTP layered on top of an encrypted TCP connection. Whilst HTTP is strictly a request/response protocol, sometimes clients need to subscribe to asynchronous notifications that are pushed from the server to the client, this has resulted in the following evolutions:
- HTTP Long Polling – is used to listen for asynchronous messages from the HTTP server, in which the client repeated initiates a GET request to wait for a notification from the server. This approach is not well suited to high rates of notifications.
- Server-Sent Events – is also based upon HTTP and relies on the server using the chunked encoding for the response. The client initiates a GET request and listens for a sequence of chunks, where each chunk holds a given notification.
- Web Sockets (WS) is an asynchronous message exchange protocol typically used with JSON based messages and initiated as protocol upgrade operation on an HTTP connection. Web Sockets Secure (WSS) is Web Sockets over an encrypted connection. Web Sockets is an effective choice for high rate notifications and provides the freedom for application developers to define their own message formats. These can be standardised with the IETF as named sub-protocols.
- CoAP (Constrained Application Protocol, RFC 7252) can be likened to HTTP over UDP and features a space efficient representation of name/value headers. CoAP is intended for use with

resource constrained devices. CBOR is a binary encoding for JSON that is often used with CoAP.

- MQTT is an OASIS protocol for topic-based routing of messages to subscribers for given topics. Routing takes place via message brokers.
- AMQP is another OASIS protocol analogous to MQTT with message-delivery guarantees such as *at-most-once*, *at-least-once*, and *exactly-once*. AMQP offers request/response as well as publish/subscribe communication patterns.

Not all IoT technologies as IP based, and as such require a gateway to the Internet. Gateways are often desirable even for IP based devices to provide greater security from cyber-attacks. This is especially important for consumer devices with weak security, e.g. devices that have the same default password for their administration, allow attackers easy access.

5.1.2.1 Challenges around configuration and discovery

One area in need of better coordination is the relationship between standards for Layer 2 and Layer 3 of the OSI model, especially in respect to auto-configuration and service discovery. IP addresses on the Internet are managed globally by the Internet Assigned Numbers Authority (IANA). IP addresses on local area networks (LANs), on the other hand, need to be managed locally, and are not globally unique. Local area IP addresses can be assigned statically, or dynamically using either DHCP or zeroconf.

Zero-configuration networking (zeroconf) involves the automatic assignment of IP addresses using a mechanism that checks that a proposed address is not already in use, e.g. using ARP or NDP. The Dynamic Host Configuration Protocol (DHCP), see e.g. RFC 2131 and RFC 8415, provides a means for clients to broadcast a request for an IP address from a service that manages the pool of available addresses for a local area network. In practice, when there are large numbers of devices connecting to a local area network, poorly configured DHCP servers can struggle to cope.

There are several competing approaches to service discovery, e.g. DNS based discovery (mDNS and DNS-SD), the Simple Service Discovery Protocol (SSDP) and the Service Location Protocol (SLP).

- SSDP is a protocol for advertisement and discovery of universal plug and play services, utilising HTTP like headers over UDP, with broadcast messages announcing the establishment or withdrawal of services. Clients can also broadcast queries for a named service, where the response is unicast. Clients can then access device services over HTTP. Each service is associated with a service-type URI and a unique service name, which are now managed by the Open Connectivity Foundation (OCF).
- DNS based discovery utilises UDP or TCP connections. Multicast DNS uses broadcast datagrams for both requests and responses, enabling clients to benefit by listening to and caching responses to queries by other clients. Services are typically associated with host names under the "local" top-level domain and include a field corresponding to a service type, which can be registered with IANA.
- SLP devices have a URL along with a set of attributes (name/value pairs). The URL scheme names are hierarchical, e.g. "service:printer:lpr". The approach allows for the use of multicast queries over UDP similar to SSDP and mDNS. You can also have directory agents that are discovered via multicast queries and can be subsequently accessed via TCP for both queries and registration of services.

The mess around service discovery is a result of competition between companies such as Apple and Microsoft, the focus on industrial consortia to promote particular solutions and a lack of broad consensus on a single shared standard approach. The Internet of Things is bringing a proliferation of service types that greatly increases the size of the challenge, and there is a big opportunity for Europe to play a leading role in respect to the coupled concerns around security, discovery and privacy as the private sector seems to have difficulties in working together on this, resulting in

unnecessary fragmentation, and increased complexity and costs for developers. A potential starting point would be for this would be an in-depth study by the AIOTI.

5.1.3 Processing Layer

This layer addresses the edge computing, data element analysis and transformation, analytics, mining, machine learning, and pervasive, considering that autonomic services are provided through ubiquitous machines in both "autonomic" and "smart" way [11].

The processing layer convert network data flows into information that is suitable for storage and higher-level processing and provides the ability to process and act upon events created by the edge devices and store the data into a database in the storage layer.

The requirements for the processing layer are connected to the need for highly scalable, column-based data storage for storing events, map-reduce for long-running batch-oriented processing of data and complex event processing for fast in-memory processing and near real-time reaction and autonomic actions based on the data and activity of devices and the interconnected systems.

Edge computing requires processing at the gateway level and the enterprise applications leverage edge devices data in end-to-end value streams involving edge devices and people within digitized processes.

Transport protocols can be used in different ways and agreement is needed to ensure interoperability. An example is the choice of media for resources transferred via HTTP. Common data formats include XML and JSON. For RDF, serialisations include XML, Turtle and JSON-LD.

Representational state transfer (REST) is a popular approach to building network APIs on top of HTTP. Different HTTP request methods are used for different purposes, e.g. GET is used to retrieve the state of a resource, PUT to upload the state of a resource or to create a new resource, and DELETE to remove a resource. The HTTP POST method is used to invoke some operation on a resource. Two approaches to describing REST APIs are: OpenAPI and the Web of Things.

There are a range of different design patterns for processing. HTTP uses the request/response pattern, Web Sockets supports asynchronous bidirectional messages where the message protocol can be specified as a Web Sockets sub-protocol. Data flow is a design pattern that is well suited for handling big data involving message streaming. Processing nodes apply their code to generate output messages when they have received all of the expected inputs. The map-reduce pattern maps large amounts of data into smaller pieces that can be processed independently, and the corresponding results are combined to get the overall result.

Another pattern is to progressively process data in stages, where each stage is a higher level of abstraction. This allows events to be inferred from noisy lower level data in combination with other data, and likewise for higher level events to be inferred from lower level ones. This can be likened to human perception, along with the means for control systems to specify the context for which kinds of features are relevant, and to direct attention to particular features as needed.

For example, consider an autonomous vehicle with a forward-facing camera. The car needs to focus attention on oncoming vehicles in case that there is insufficient room for both to pass each other on a narrow city street. In this situation the car could look for spaces between parked cars on the side of the street to allow it to pull in to let the other car pass by safely, or vice versa.

5.1.4 Storage Layer

The IoT stakeholders addressing this layer consider the efficient storage and organization of data and the continuous update of data with new information, as it made available through the capturing and processing channels [11]. Archiving the raw and processed data addresses the offline long-term storage of data that is not needed for the IoT system's real-time operations. Centralized

storage considers the deployment of storage structures that adapt to the various data types and the frequency of data capture.

Relational database management systems can be used that involves the organization of data into a table schema with predefined interrelationships and metadata for efficient retrieval for later use and processing. Storage technologies such NoSQL key-value stores are used to support big data storage with no reliance on relational schema or strong consistency requirements typical of relational database systems. For autonomous IoT applications, the storage can be decentralized, and data is kept at the edge or at the objects that generate it and is not sent up the system.

Data storage is subject to multiple considerations:

- Local buffering for block transfer as a means to extend battery life and make more efficient use of IoT communication technologies.
- Local storage when processing cannot tolerate the latency involved in remote access to a data centre.
- Local processing can avoid the need to store large amounts of data, e.g. consider a traffic camera observing traffic passing through a junction. Rather than storing the video, it is sufficient to just store the number of vehicles passing through in each time period.
- Federated processing that avoids the need for central storage of data. An example is federated machine learning where personal data never leaves the owner's devices, as each device works to incrementally improve the models which are then consolidated centrally and distributed to where they are needed.
- Security – keeping data out of reach of attackers.
- Privacy – fulfilling the obligations of GDPR and other related regulations, e.g. in respect to enabling data owners to access and request the deletion of personal data.
- Dependability – using redundant copies as a means to defend against faults, fire and floods.
- Throughput – being able to support the expected peak demand for data access.
- Duration – how far back in time will data be needed, and thus how much data needs to be stored? Stream processing of data can reduce the need to keep old data around.
- Transactional consistency – maintaining consistency across data via the means to roll back aborted transactional updates.

Some companies are providing solutions that integrate storage and processing enabling application code to run efficiently within the storage module, so that large volumes of data can be processed in parallel at very low latency, e.g. for deep learning algorithms, data mining and data analytics. Cloud vendors offer flexible storage solutions according to need, and drawing upon their experience with operating global services, e.g. Web search and online marketplaces.

5.1.5 Abstraction Layer

This layer provides the interfaces and the event and action management through rules engine to allow mapping of low-level sensor events to high-level events and actions, while assuring the basic analytics for data normalization, reformatting, cleansing and simple statistics [11]. IoT systems scale corporate and global level and require multiple storage systems to accommodate IoT device data and data from traditional Enterprise Resource Planning (ERP), Human Resources Management System (HRMS), Customer Relationship Management (CRM), and other systems.

The data abstraction functions are rendering data and its storage in ways that enable developing simpler, performance-enhanced applications. Data abstraction layer processes data and reconcile multiple data formats from different sources, assuring consistent semantics of data across sources, confirming that data is complete to the higher-level application, consolidating data, providing access to multiple data stores through data virtualization, normalizing or de-normalizing and indexing data to provide fast application access, protecting data with appropriate authentication and authorization.

The IoT is highly fragmented with a huge number of different technologies and standards, and a general lack of interoperability. The Abstraction Layer seeks to counter this fragmentation by presenting a unified framework that hides the underlying complexity from application developers.

One such approach is W3C's Web of Things. This models physical devices as digital twins:

- Virtual digital objects that stand for physical and abstract entities
  - *Sensors, actuators, heterogeneous information services,*
- that are exposed to client applications as local software objects
  - *Clients can interact with the object's properties, actions, and events*
  - *Client applications do not see or need to deal with HTTP, Bluetooth, etc.*
    - *those details are handled by the web of things client platform*
- and used as part of semantic descriptions
  - *The kind of sensor, its physical location, units of measure, …*
  - *Object histories that can be used to monitor and assess changes over time*

W3C has been working on the Web of Things for several years and has recently published proposed Recommendations (W3C's term for its standards) for thing descriptions using JSON-LD, and on architectural considerations for the Web of Things. Supplementary notes cover security considerations and a proposed scripting API.

- Web of Things: Architecture
- Web of Things: Thing Descriptions
- Web of Things: Scripting API
- Web of Things: Binding Templates
- Web of Things: Security and Privacy Guidelines
- Web of Things: Current Practices

Thing Descriptions include: a programming language neutral description of the object API exposed to client applications, communications and security metadata for use by the client platform to communicate with the server platform that exposes the things, and semantic descriptions of the kind of thing, what it measures, the units of measure, the sensor location, the operations it supports, and so forth.

W3C's framework is based on the Resource Description Framework (RDF) and the JSON-LD serialisation of RDF. This can be contrasted with the context information management framework developed for FIWARE (FIWARE-NGSI v2) which is based upon Property Graphs and exposed to applications via a REST API. This has now been standardised at ETSI as the NGSI-LD API (ETSI GS CIM 009), drawing upon earlier work by OMG. See section 8.3 of this report for an account of the evolution of database technologies, and the relationship between RDF and Property Graphs.

Some related work includes the oneM2M mechanism for resource discovery (TR 0057) in which a client initiates a RETRIEVE operation on a server, involving a resource path and optional search filter expressed as a conjunction of attributes, e.g. to select resources created after a given date and time. W3C is working on a discovery interface for the Web of Things implemented as a distributed application involving clients, servers and directory services. The discovery interface supports local discovery for Things exposed on the same device as the client, remote discovery based upon querying a Thing Directory, or discovery using a supported multicast protocol.

The concept of a digital twin that applications can interact with as a proxy for the physical device, has its limitations. In many cases, the emphasis is on feeding data into a centralised framework for services. Applications interact with this data and have no need for remote interaction with the IoT devices themselves. This allows for a simpler architecture in which "connectors" gather data from IoT devices and feed it into a cloud-based system along with the metadata that describes it, e.g. data models and ontologies. Applications focus on the use of the data and metadata, e.g. in the

healthcare domain, to monitor a patient's condition over time and to support medical staff in selecting appropriate interventions.

5.1.6 Service Layer

This layer integrates the middleware that sits on top of networks [11]. The IoT device streams that provides data management and data analytics are vital functions in IoT systems where large amounts of sensor generated data and events must be logged, stored and processed to generate new insights or events on which business decisions can be taken.
Services typically perform a single operation or closely related set of operations. Services are most often accessed by other programs and are designed to target part of a large problem domain. For an IoT marketplace, service providers register services for use by service consumers. Services may themselves act as users of other services where a problem can be decomposed into smaller problems addressed by existing services. "Microservices" are fine grained and address a narrowly defined task.
More generally, service-oriented architecture (SOA) is where services are provided to other software components via a communication protocol over a network, e.g. a REST API over HTTP. Each service:

- logically represents some task with a specific outcome
- is self-contained, is a black box for its consumers in the sense that they do not need to be aware of the service's inner workings
- it may be composed from other underlying services
- is discoverable on the basis of associated metadata that describes the function of the service, and the terms and conditions for its use.

5.1.7 Application Layer

This layer is offering the software platforms that are suited to deliver the key components for implementing various IoT applications that are connecting users, business partners, devices, machines, and enterprise systems with each other and the information interpretation is provided [11]. Software at this layer interacts with the service layer, while the software applications are based on vertical markets, the nature of device data, and business needs. At this layer, many applications are addressed such as mission-critical business applications, ERP, specialized industry solutions, mobile applications, analytic applications that interpret data for business decisions, etc.

Applications perform a wide range of operations and may even expose some of these as services. Applications are typically accessed by human users and target a whole problem domain. Applications are implemented on top of software APIs or services. Applications may need to be installed, e.g. as is the case for native applications on smart phones, tablets and desktop computers, or they can be Web based and accessed through a browser. Web applications can be designed to adapt the user experience according to the class of device, e.g. smart phone vs a desktop with a much larger display. Applications should be designed for accessibility by all people regardless of disabilities or severity of impairments.

Web applications and many native applications involve a mix of local and remote code accessed via a fast network connection, e.g. Wi-Fi or 4G. It thus makes sense to design IoT platforms around a data store and an HTTP server. The data store is likely to include structured and unstructured data, e.g. graph data, HTML, images, style sheets and scripts. It is increasingly common for a front-end application server to have many back-end servers designed for handling different tasks.

Native applications are typically discoverable and installable from an application store, e.g. Google Play for devices using the Android operating system. Web applications can be discovered

using a Web search engine, e.g. Microsoft's Bing, and do not need any installation, although users can create bookmarks for easier access on subsequent occasions.

5.1.8 Collaboration and Processes Layer

This layer includes enterprise systems and the exchange of data among platforms [11]. The layer also addresses the processes that involves people, organisations that use applications and associated data for their specific needs or for a range of different purposes, to provide the right data, at the right time, to perform the right thing.

End to end security is addressed for each layer and as the data is moved across the layers to secure each device or system, provide security for all processes at each level, secure end to end exchange and communication between each layer.

This layer is about people, business processes and business objectives and usually transcends multiple applications. To satisfy business objectives, organisations use data governance to define how data is accessed and treated within a broader data manage strategy, which itself is concerned with the implementation of architectures, tools and processes to achieve stated data governance objectives. As such both data governance and data management are coupled to an organisation's IoT strategy.

For ecosystems of services based upon the IoT, either business to business, or business to consumer, there is a need for clearly defined terms and conditions, including attention to data sovereignty, i.e. who owns the data, what restrictions are placed on use of the data and so forth. Standardised vocabularies for this would be helpful in respect to searchability.

5.2 Standards for the Cross-Cutting Functions Dimension

5.2.1 Identifiability

Identifiability enables an IoT based system to identify unique devices and exchange the necessary information according to a given application. The devices can vary from a few small sensors/actuators to a system that interconnects a huge number of devices with a capacity to deliver complex AI based services (e.g. Big Data). In this context it is important to follow the General Data Protection Regulation (GDPR) regarding personally identifiable information.

Unique identifiers can create challenges for privacy. One mitigation is to use pseudonymous identifiers that are only applicable in a limited narrow context. Another mitigation is limit access to the identifiers, e.g. firstly, by using encrypted communication channels to block snoopers, and secondly, via access control that constrains discovery to authorised clients.

Another challenge is in respect to ownership and management of identifiers. Identifiers based upon the Domain Name System are at risk when the domain name ownership is transferred from one party to another, e.g. when a company goes out of business, or merges or is taken over by another. Long lasting identifiers are only as good as the organisation that manages them.

Of particular note, digital object identifiers (DOI) are persistent identifiers that have been standardised by the ISO and supported by the International DOI Foundation (IDF). The aim is to maintain up-to-date mappings from DOI to URLs for the intended resource. DOI can be contrasted with other kinds of persistent identifiers such as ISBNs and ISRCs which only aim to identify their referents uniquely.

Most identifiers are managed centrally. By contrast, decentralised identifiers are a new class of identifier that are globally unique, resolvable with high availability and cryptographically verifiable. Such identifiers do not require a centralized registration authority because they are registered using distributed ledger technology or other forms of decentralised networks. For more information see W3C's Distributed Identifiers (DIDs).

An identifier identifies single entity or class of entities within a specific context. Identifiers are used for different purposes in IoT applications, and AIOTI WG03 (IoT standardisation) classify the identifiers in IoT as follows [25]:

- Thing identifier - identify the entity of interest of the IoT application like physical objects or digital data. Basically, be anything that one can interact with. Typical applications are predictive maintenance, asset tracking, and quality control.
  Numerous standards are available for identifying things. They are often defined for specific domains or specific types of entities, but some are used in several domains and for different types and classes of entities. Some standards provide mechanisms to enable multiple identification schemes to interwork in the same IoT application.
- Application and service identifier - identify SW applications and services. This also includes identifiers for methods on how to interact with the application or service. A typical example is IoT platform services.
  Application and service identifiers are usually defined in the context of the specific platforms (e.g. service platform, operating system) on which they are provided. This can be based on open standards or be proprietary. In case the platform is standardized also the application and service identifiers are standardized.
- Communication identifier - identify communication points and communication sessions. Examples on usage are: (*i*) Low Power Wide Area Networks (LPWANs) defined by ETSI GS LTN 002 using uniquely assigned communication identifiers to identify end devices etc., (*ii*) In Ethernet Networks (IEEE 802.3) the Media Access Control (MAC) address is an identifier for communication endpoints at the data link layer, usually assigned by the manufacturer, and (*iii*) IP addresses (IPv4, IETF RFC 791 and IPv6, IETF RFC 429) are used in IP networks to identify communication endpoints at the network layer.
  Communication identifiers are essential for a communication protocol and impact its functionality like routing and switching. Usually the identifier scheme cannot be changed without major changes to the protocol itself. Identifiers are therefore defined as part of the specific communication protocol standards.
- User identifier - identify users of IoT applications and services. Examples on users are humans, legal entities or SW applications that access and interact with the IoT application or service.
  User identifier formats are usually defined by the specific system for which user access is needed. They could be provided by the users and checked by the system for uniqueness or assigned by the system.
- Data identifier - identify both specific data instances and data types. Examples on usage are: Digital twins' representation, time series data set like sensor data provided in time intervals, and property types like weight, dimension, and temperature characteristics.
  Various standards for the identification of data sets, files, streams, metadata, data types and other data elements exists. Some standardized solutions provide support for multiple identification schemes in order to cover already existing schemes and enable the definition of domain and context specific schemes.
- Location identifier - identify location within a geographic area. It could vary from geospatial coordinates to postal addresses, or even room numbers. Examples on usage are: Gods tracking, and real estate maintenance.
  Location identification is important in many IoT applications. Location identification standards exist for the objective naming of a geographical location. This information is often carried out in IoT application to keep track of where an event happens or where things are supposed to be or should go to.
- Protocol identifier - identify protocol essential information. The identifiers inform for example communication protocols about the upper layer protocol they are transporting or applications about the protocol they must use in order to establish a specific communication exchange.

Like communication identifiers, protocol identifiers are usually defined as part of the protocol that uses them.

5.2.2 Trustworthiness

Trust is defined has having confidence, faith or hope in someone or something. Trust is established either through experience of how the target person or thing behaves, or through the declaration by a third party who is trusted when they say that the target person or thing is trustworthy. In general trust is not associative nor commutative. If "a" trusts "b" and "b" trusts "c" then "a" will not necessarily trust "c". Likewise, just because "a" trusts "b" does not necessarily mean that "b" trusts "a". Moreover, trust is contextualised and a matter of degree rather than true or false. People can be trusted to do some things, but not necessarily other things. Trust is thus a very human concept and difficult for machines to reason with.

Trust in an IoT system must be bootstrapped, for example, using secure boot from trusted tamper proof hardware modules that allows a system to verify the integrity of its hardware and software. This process is built on trust in the identity of a given party (e.g. the device manufacturer) in terms of cryptographic operations. To retain trust, the software on IoT devices should be securely updatable to fix security flaws as they are discovered. If that is impractical, the IoT devices and the communications pathways need to be physically protected from attackers.

A further challenge is how to transfer trust when either the owner of a device changes, or when the organisation managing the device changes, e.g. when a company goes bust or is taken over by another. Devices need to be designed for easy installation, and similarly easy to reset when transferring ownership, e.g. when a parent hands a device over to a child or sells it to another person on eBay. The reset process needs to clear any personal information on the device.

Trust further depends on limiting access to authenticated devices or services which are uniquely identified to participate in the decision-making processes of a system [24]. This makes it possible to report the source of vulnerabilities and inconsistencies.

Trustworthiness becomes an essential requirement, since an increasing amount of AI-enabled systems become connected through the IoT, trustworthiness becomes an indispensable requirement. Trustworthiness will become multi-dimensional, far beyond verifying identity. Consequently, trust will no longer be true or false, but rather about degrees of trustworthiness that will control the access levels of devices and users to critical services and systems.

Trustworthiness enables IoT systems to be trusted, by only allowing authenticated devices or services which are uniquely identified to participate in the decision-making processes of a system [24]. This makes it possible to report the source of vulnerabilities and inconsistencies. Trustworthiness becomes an essential requirement, since an increasing amount of AI-enabled systems become connected through the IoT, trustworthiness becomes an indispensable requirement. Trustworthiness will become multi-dimensional, far beyond verifying identity. Consequently, trust will no longer be true or false, but rather about degrees of trustworthiness that will control the access levels of devices and users to critical services and systems.

Trustworthiness in IoT applications, and IIoT solutions in particular, are subjected to compliance requirements and standards [26]. The requirements are often domain specific and driven by trade and governmental regulations. IIoT systems should be designed and deployed in accordance with relevant requirements and standards. According to the Industrial Internet Consortium some mainstream and minimum examples of IIoT compliance requirements are [26]: *(i)* General Data Protection Regulations (GDPR), *(ii)* Workplace health and safety regulations, *(iii)* Maintenance of records for auditing purposes, and *(iv)* The Health Insurance Portability and Accountability Act (HIPAA).

Five categories within trustworthiness domain together with related domains and its potential consequential events are listed below [26]. It is important to have in mind that this framework is "heavily" dependent of the IIoT system of interest.

- Personal and community welfare - Related to characteristics like security, safety, reliability, resilience, and privacy. Can lead to environmental consequences, reputational loss, personal injury or even loss of life.
- Data loss - Related to characteristics like security and privacy. Can lead to consequences like compromise of personal data, identify theft and compromise of commercial sensitive data.
- System - Related to characteristics like security, safety, reliability, and resilience. Can lead to consequences like damage to physical systems, reduce capacity to operate, compromised data records, and process failures.
- Commercial - Related to characteristics like security, safety, reliability, resilience, and privacy. Can lead to consequences like negative brand impact, reduces revenue, increased costs, and loss of customers.
- Legal - Related to characteristics like security, safety, reliability, resilience, and privacy. Can lead to consequences like non-compliance with prevailing regulations and standards.

The Technical Committee ISO/IEC JTC1 SC41 on Internet of Things and related technologies has the following standards under development regarding "Trustworthiness" [27]:

- ISO/IEC AWI 30147: Information technology - Internet of Things - Methodology for trustworthiness of IoT system/service. The status of this work is "Proposal" (i.e. new project approved).
- ISO/IEC AWI 30149: Internet of Things (IoT) - Trustworthiness framework. The status of this work is "Proposal" (i.e. new project approved).

In addition, the Technical Committee ISO/IEC JTC1 SC42 on Artificial intelligence has the following standard under development regarding "Trustworthiness" [27]:

- ISO/IEC PRF TR 24028: Information technology - Artificial intelligence - Overview of trustworthiness in artificial intelligence. The status of this work is "Approval" (i.e. proof sent to secretariat or FDIS ballot initiated).

5.2.3 Security

In respect to the IoT, security is about the protection of devices, computer systems and networks from theft, damage, disruption, or misdirection of the service they provide [12]. Some IoT devices and technologies lack strong security. This can be addressed through a combination of physical security (e.g. preventing access to the devices and cabling) and cybersecurity for gateways that connect the IoT device to untrusted networks such as the Internet.

Security best practices include the requirement for being able to securely update the software and firmware used by a particular device to fix vulnerabilities. This may not be feasible for very resource constrained devices, necessitating the use of security gateways to protect them from external attackers. To prevent snooping and man-in-the middle attacks, the communication channels need to be protected end to end, which is vital to ensure robustness against all types of attacks into the IoT system and its constituents [24].

Software access, including across the network, needs to be limited to authorised clients. Access control policies can take a number of factors into consideration, e.g. the identity of the client, the resource to be accessed, the time of day, etc. A simple approach is to use access control lists, in which clients are associated with roles, and roles with authorisation for particular services/resources. A more powerful approach is the OASIS extended access control mark-up language (XACML) which provides a fine grained, attribute-based access control policy language.

A contrasting approach is to use capability tokens that give the bearer the right to access some service or resource. The bearer token is time limited and issued after strong authentication of the

client. JSON Web Tokens are an open industry standard for bearer tokens. A related approach is taken by OAuth which provides a means for a user to grant an application, such as a website, time limited access to personal data held by another site.

Security management deals with the disciplines involved in managing identities, trust, vulnerabilities and threats. Best practices start by trying to keep attackers out. If that fails, then how to detect and mitigate attacks, both in the short and long term. Security resiliency is about limiting the damage that an attacker can do, for example, by dividing a system up in to separate security domains where the boundaries between domains acts as barriers that the attackers will have to overcome if they are to achieve their goals. Systems should be designed to detect intrusions and unauthorised or unexpected behaviours (e.g. as spotted through machine learning).

The security system can be designed to respond automatically with the first level of mitigations whilst summoning human cybersecurity staff to the rescue. An example is where a security monitor detects excessive requests to a given resource in a short period of time (e.g. denial of service attacks), which can be mitigated by immediately closing the connection at the firewall before the connection is made with the internal server. Systems may include so called "honeypots" that are designed to divert the attacker's attention and facilitate the erection of barriers around the infection (*aka* "islanding") allowing services to continue operating as normal.

For further reading:

- European Union Agency for Network and Information Security (ENISA) contributes to securing Europe's information society by raising awareness and by developing and promoting a culture of network and information security in society thus contributing to the proper functioning of the internal market.
  - ENISA IoT Security Standards Gap Analysis
- European Data Protection Board (EDPB) is an independent European body, which contributes to the consistent application of data protection rules throughout the European Union, and promotes cooperation between the EU's data protection authorities.
- IoT Security Foundation is a collaborative, non-profit, international response to the complex challenges posed by cybersecurity in the expansive hyper-connected IoT world.
- Open Web Application Security Project (OWASP) is a non-profit foundation that works to improve the security of software to secure the Web.
- ETSI TS 103 645 covers cybersecurity for consumer devices connected to network infrastructures and provides thirteen recommendations:
  1. No universal default passwords
  2. Implement a means to manage vulnerability reports
  3. Keep software updated
  4. Securely store credentials and security-sensitive data
  5. Communicate securely
  6. Minimize exposed attack surfaces
  7. Ensure software integrity
  8. Ensure that personal data is protected
  9. Make systems resilient to outages
  10. Examine telemetry data
  11. Make it easy for consumers to delete personal data
  12. Make installation and maintenance easy
  13. Validate input data
- The UK has introduced new legislation to improve security standards of internet connected household devices, that requires devices to have unique passwords with no factory reset option; manufacturers will be required to create reporting functions for vulnerabilities, and at the point of sale must inform consumers of the minimum length of time that security updates will be provided.

5.2.4 Safety

Safety is the ability of the IoT system to operate without harmful states and catastrophic failures [12]. Safety enables systems to protect persons and objects during use and operation [24]. IoT/AI based systems that operate physically collaboratively with or next to humans through robots (or other machines) must not exhibit random or unpredictable behaviour. Safety by design is essential, in compliance with relevant safety standards. The employed IoT/AI systems must be robust against unwanted data and operate with extremely low latency to react to unforeseen events quickly and appropriately.

Consumer electronic devices are subject to strict safety regulations, e.g. physical safety such as protection from electrical shocks, the risks from devices initiating fires, or the use of hazardous materials. Non-physical concerns include electromagnetic interference, and the harm that be done through theft, threatened or actual disclosure of personal data, or through fraud.

The General Product Safety Directive (GSPD) 2001/95/EC defines EU rules for product safety: products are deemed safe if they comply with all statutory safety requirements under European or national law.

5.2.5 Privacy

Privacy enables IoT/AI based systems that operate on mission- and business-critical data [24]. This implies both limiting access to and placing restrictions on certain types of data with the objective of preventing unauthorized access as well as protecting data from being modified or corrupted without detection. It is an advantage that such data are processed locally at the edge and only leverage data available within privacy limits.

Privacy is about enabling an individual person or group to keep chosen information about themselves to themselves, i.e. to keep it private. The European GDPR regulations gives individuals the right to access, and under certain conditions, the right to have deleted, the personal information about them, that is held by an organisation. Organisations need to secure personal information and limit the operations that can be performed on it.

This implies both limiting access to and placing restrictions on certain types of data with the objective of preventing unauthorized access as well as protecting data from being modified or corrupted without detection. It is an advantage that such data are processed locally at the edge and only leverage data available within privacy limits.

Data anonymisation is the process whereby personal identifying information is removed prior to processing information, for example when applying machine learning across data collected from many people. Data should only be made available under binding conditions that precludes de-anonymisation via combining multiple sources of information.

Privacy and trust are coupled. It is common for businesses to require credentials as a precondition for doing business with them. This represents a loss of privacy in addition to the irritation involved in having to type in or otherwise provide the requested information. There are opportunities to exploit technical means to minimise this, for example to use zero-knowledge proofs that a user is in possession of a credential without having to provide access to that credential. For instance, to prove that a person resides within a given city without having to disclose his or her address, and likewise, to prove that someone is eighteen years and older without disclosing their birthdate. Such proof relies on strong credentials that are widely trusted on the basis of the processes used for issuing them.

The IoT is an Orwellian threat to privacy on account of the many sensors that people carry on them (e.g. in their smart phone) or are exposed to in their homes, or when they are away from home. This threat is magnified by the level of tracking that is commonplace online in the name of

analytics and tailored advertising. The continuing acceptance of the IoT depends on users being able to trust the organisations that operate them.

5.2.6 Connectivity

Connect ability is the ability of the IoT system to connect securely, anytime, and anywhere to any available network [12]. Connectivity in IoT/AI based systems enables devices, programs, computers and/or systems to communicate with each other. The number of connected IoT devices data are increasing rapidly, including both cloud and edge computing. Faultlessness, low latency communication and end-to-end security are essential properties for the underlying connectivity system.

5.2.7 Resilience

Resilience is the ability of the IoT system to transform, renew, resist, respond and recover timely from damaging effects or states [12]. Resilience enables IoT to "always" operate in stable states and return to stable states if potential failures occurs [24]. Resilience is especially important for safe support in our digital economy and should be able to detect failure and initiate measures for mitigation.

Resilience more generally covers cyber-attacks, hardware and software faults, and spikes in demand. Resilient systems are those that are design to be able to continue to operate as best they can when individual subsystems have been compromised. This can be contrasted with systems that collapse as the failure of one subsystem leads to the failures of others in a repeating wave that brings the whole system down (the so called “domino effect”). In particular, this means that systems should be designed to avoid undue reliance on a centralised component that becomes the Achilles heel in a system that is otherwise strong overall.

5.2.8 Reliability

Reliability is the ability of the IoT system to deliver and accomplish services as specified within given constraints [12]. Reliability enables IoT/AI based systems to operate without failures, outages and regular human intervention [24]. Reliability is essential for productivity and is a key prerequisite for IoT/AI systems in continuous operation with short maintenance time in mission-critical operations. Reliability may be associated with quality of service commitments and is related to design for resilience. IoT services based upon wireless communications may in some cases suffer from adverse weather, electrical noise, and interference from other wireless traffic. For cellular networks, another factor is poor coverage, e.g. in rural areas. A reliable and robust communication network is key to the future of autonomous vehicles.

5.3 Standards for the System Properties Dimension

The third dimension in the 3D Reference Architecture Model is represented by the “System Properties” to support the discussion about the expected system properties of the IoT system between different involved parties (e.g., users, contractors, designers) and identify the elements in support (e.g., functional building blocks, APIs) and those missing.

To some extent, the notion of system property has a more “open” character than the “Layers” and “Cross-cutting Functions” dimensions. Given the nature of the IoT system under consideration (e.g., business domain, integration with legacy), the list of “system properties” could be defined on a more “IoT project-oriented” basis than the other two dimensions. In addition, it should be noted that there is no commonly agreed definition for some of the “system properties” analysed.

Secondly, the “system properties” dimension is not only directed to the identification of applicable standards: the discussion between stakeholders may also lead to the identification of applicable codes of conduct or policy directives.

The definition of the "system properties" is developed in the companion deliverable D06.03. Regarding standards, some points may have to be noticed that are addressed in Table 7. For more information, we refer to CREATE-IoT deliverable D06.03.

Table 7: Some issues to consider regarding the Properties dimension

Property	Points of notice
Interoperability	Interoperability is a characteristic of a product or system, whose interfaces are perfectly able, to work with other products or systems, at present or future, in either implementation or access, without any restrictions [28]. More specifically, interoperability is defined as the degree to which two or more IoT systems/platforms, can exchange information/knowledge and use the information/knowledge that has been exchanged.
Composability	Composability is defined as an IoT system property that address the inter-relationships of components, with composable IoT systems integrating components that can be selected and assembled in various combinations to satisfy specific user requirements. In IoT the features required for a composable IoT system is to be self-contained (modular) and can be deployed independently and to be stateless and can treat each data request as an independent transaction, unrelated to any previous requests.
Scalability	Scalability is the ability of a process, network, software, or organization to grow and manage increased demand [28]. Scalability is often a sign of stability and competitiveness. Scaling can refer to the ability to support increasing numbers of devices, e.g. millions of cars or billions of smart phones. It can also refer to the amount of data and its throughput.
Integrability	Integrability is defined as the degree of effectiveness and efficiency with which an IoT system can be successfully integrated within heterogeneous IoT systems/platforms and sub-systems or other types of systems including platforms. This can be contrasted with composability of IoT services.
Manageability	Manageability is the ability to manage the IoT system to ensure continuous operation. This is important to being able to handle very large numbers of IoT devices in a convenient and cost-effective manner. That includes, for instance, the installation of new devices, software upgrades to fix vulnerabilities, and the transfer of ownership of individual devices, and re-establishment of trust.
Dependability	Dependability is the ability to deliver a service that can justifiably be trusted [12]. Another definition of dependability is the ability to avoid service failures that are more frequent and more severe than is acceptable. This relates to quality of service commitments, and the speed with which vulnerabilities can be fixed by rolling out software updates.
Availability	Availability is defined as the degree to which an IoT system/platform is operational and accessible when required for use. The ability of the IoT system to deliver services and information when requested [12].
Intelligence	Intelligence in the meaning of Artificial Intelligence (AI) and AI features, algorithms, techniques, and methods used in different IoT layers of a system to provide different levels of intelligent functions and behaviours. AI is also related to machine learning, including deep learning based upon multi-layer artificial neural networks, which can be contrasted with techniques based upon reasoning over symbolic representations.

6 IoT and Data

6.1 The Importance of Data

After ushering in the forefront, being developed in support of more and more complex services and deployed with massive numbers of associated devices, the IoT systems have started to produce enormous amounts of data. In particular, that data is essential for ecosystems of services, e.g., smart cities, healthcare, manufacturing, transport, etc.

6.1.1 The European Data Strategy

Taking into account that profound strategical change and the opportunity it creates for Europe to leapfrog on data-oriented business models, the EU promotes a data strategy as a major aspect of the EU Digital Single Market and emphasises the major role of data spaces. To quote from the European Strategy for Data:

- Data is an essential resource for economic growth, competitiveness, innovation, job creation and societal progress in general.
- Data driven applications will benefit citizens and businesses in many ways. They can:
  - Improve health care
  - Create safer and cleaner transport systems
  - Generate new products and services
  - Reduce the costs of public services
  - Improve the sustainability and energy efficiency
- Businesses will have more data available to innovate. This will be done by launching practical, fair and clear rules on data access and use, which comply with European values and rules such as personal data protection.

The European strategy for data aims at creating a single market for data that will ensure Europe's global competitiveness and data sovereignty. Common *European data spaces* will ensure that more data becomes available for use in the economy and society, while keeping companies and individuals who generate the data in control.

- The EU will create a single market for data where:
  - Data can flow within the EU and across sectors, for the benefit of all.
  - European rules, in particular privacy and data protection, as well as competition law, are fully respected.
  - The rules for access and use of data are fair, practical and clear.
- The EU will become an attractive, secure, and dynamic data economy by:
  - Setting clear and fair rules on access and re-use of data.
  - Investing in next generation standards, tools, and infrastructures to store and process data.
  - Joining forces in European cloud capacity.
  - Pooling European data in key sectors in EU-wide common and interoperable data spaces.
  - Giving users rights, tools, and skills to stay in full control of their data.

The data economy involves the provision and consumption of data services. The European Commission envisages an open marketplace based upon the FAIR principles that require data to be: *Findable* via rich searchable metadata; *Accessible* via standard protocols; *Interoperable* via use of shared vocabularies; and *Reusable* via clear and accessible data usage licenses. See the European Commission's paper on Turning FAIR into reality.

In more detail, the interoperability requirement calls for the use of a formal, accessible, shared and broadly applicable language for knowledge representation.

6.1.2 The role of technologies and standards in support of the EU strategy

The term "data space" can be defined as collections of heterogeneous data without the need for prior agreement on semantic integration. For we need:

- Data should be accompanied with metadata that allows for incremental work on integration across data sources.
- Data and metadata should use a formal, accessible, shared and broadly applicable language for knowledge representation.

With the above definition data spaces can be considered as containers for *knowledge graphs*, i.e. data with their associated data models.

Acknowledging the major role of data spaces goes together with putting the accent on how to master the associated technologies such as graph databases and knowledge graphs (data and models), as well as statistics, data sovereignty, and so forth.

The way that people think about the IoT depends on their background and has an impact on their ability to address data in the IoT service lifecycle. Communications engineers naturally think about IoT in terms of communication technologies, focusing on the lower layers and the associated technologies, e.g. Bluetooth and 5G, along with the data transport protocols such as CoAP and MQTT. In contrast, developers of services are thinking about the IoT in terms of data and the value it can bring to a service, focusing on the data and how it can be used to support the end-user's needs.

Since much of the value of the IoT is in the data and the services it supports, the means to combine heterogeneous information sources become especially important. For the IoT, this points to opportunities to focus on a *common federated framework* for representing, accessing and manipulating information (data and metadata), and how this relates to facilitating open ecosystems of services based upon the digital single market.

The following sections consider current and likely future trends in respect to *standards*. This includes some of the work done on data spaces in the LSPs (that is also developed in the companion deliverable D06.03 [3]).

6.2 Technology enablers for data spaces

Two important technology developments are addressed in this section regarding the possibility to improve the effectiveness of data management in the context of IoT architectures. Firstly, the emergence of Edge Computing as a support of an efficient management of data as close as possible to where it is produced (i.e. on the edge of the network, close to the devices). Secondly, the advent of a new type of databases, namely graph databases, taking advantage of the progress of data base technology in order to provide a more efficient management of semantic information.

6.2.1 Edge Computing and Peer to Peer Services

Data storage and processing for the IoT can usefully take place at different locations according to need:

- In the IoT devices hosting sensors and actuators at the network edge
- In nearby hubs and gateway devices
- In cloud-based systems at different scales from individual servers to server farms

Processing data at or close to the network edge can reduce the load on cloud servers. Such processing can generate higher level interpretations, as well as mitigating gaps in the data caused by faulty sensors or problems caused by electrical noise. Edge computing can be defined as a distributed computing paradigm that brings computation and data storage closer to the location where it is needed. Some advantages over centralised cloud systems include:

- Reduced latency in respect to computer games and cyber-physical control systems
- Faster data transfer rates through high speed local connections
- Better control over data ownership when data is stored on the user's premises

The Edge Computing Consortium Europe (ECCE) was founded in 2019 to specify a reference architecture for edge computing, refence technology stacks, to identify gaps for further work, recommendations for best practices, for synchronisation with related initiatives and standards development organisations, and for promotion of results. Founding members include Huawei, Analog Devices, Arm, Bombardier, B&R Automation, Fraunhofer Institute for Open Communication Systems (FOKUS), German Edge Cloud (GEC), German Research Center for Artificial Intelligence (DFKI), HARTING IT, IBM, Intel, KUKA, National Instruments, Renesas Electronics, Schneider Electric, Software AG, Spirent, and TTTech.

Edge computing is related to peer to peer architectures for communication services, where a set of nodes need to communicate with each other, e.g. for multiplayer computer games or peer to peer video conferencing. If each node were to communicate directly with each other the number of connections increases as the square of the number of participants. A more efficient architecture is to form the nodes into a scale free network where nodes forward messages on behalf of others in a friend of a friend network, where some nodes have more friends than others. This introduces a routing delay that scales very efficiently with the number of participants, see Cohen and Havlin 2003.

Algorithms have been developed to dynamically form nodes into scale free networks with guarantees on stability of convergence, based upon node processing power, storage capacity, uplink and down link bandwidth, see Ponec et al. 2009. Such algorithms are resilient in respect to changing conditions including device and localised network outages,

6.2.2 The Evolution of Database Technologies and the role of AI

As the amount of data increases it is common to take advantage of database management systems (DBMS) that support efficient organisation and access to related data. DBMS typically provide support for data models, data update, data retrieval and administration, including user accounts, backups and restoration.

6.2.2.1 Enabling technology evolutions

The evolution of databases has been driven by technological improvements in storage media and computer power. Early systems involved sequential access to secondary storage based upon magnetic tape and punched cards. Magnetic drums and disks appeared in the mid-1960's and enabled random access. Early DBMS (e.g. CODASYL) provided access to data via navigating between data records. Hash tables and B-trees enabled key-based data indexes. CODASYL is complex and gave way to table-based relational databases (RDBMS) that were more space efficient and easier to program. RDBMS became the dominant form of database management systems by the mid 1980's along with the SQL query language.

The emergence of object-oriented computing in the 1990's gave rise to the development of object databases as a means to avoid the cost of translating between objects and RDBMS tables. In the 2000's XML databases appeared for access to structured documents, reflecting a strictly hierarchical approach to data. NoSQL databases offer high performance for indexed data where related data is grouped together to avoid the cost of the "join" operations in RDBMS.

6.2.2.2 Graph databases, property graphs

Graph databases involve graphs of vertices, edges, and attributes. Commercial graph databases appeared in the 2010's, e.g. Neo4J and Oracle Spatial and Graph, and in recent years have been followed by many others. Graph databases reflect conceptual relationships expressed in terms of

labelled directed edges between vertices, as pioneered in the second half of the twentieth century with work on Semantic Networks, e.g. Ross Quillian's Ph.D. thesis (1968) and Anderson and Bower's work on modelling human memory (1973), and ideas that go back over two thousand years to Aristotle.

Starting in the late 1990's, W3C has developed a suite of standards for graphs based on the Resource Description Framework (RDF). RDF graphs use URIs for vertices and edge labels. These URIs may be dereferenceable to obtain further information, enabling a web of data. In addition, URIs as globally unique identifiers provide the basis for standardised vocabularies. W3C's standards for RDF Schema and OWL (the Web ontology language) have been widely adopted. W3C also developed SPARQL as a query language for RDF, and SHACL as a graph constraint language. There are a variety of standards for serialising RDF: RDF/XML, Turtle and most recently JSON-LD, which builds upon the popularity of JSON (JavaScript Object Notation) for data exchange.

RDF focuses on binary directed relationships (<subject, predicate, object> *aka* "triple") and is awkward when it comes to annotating relationships, which necessitates the use of a mechanism called "reification" in which the triple is modelled as a vertex with separate relationships for the subject, predicate and object. Commercial applications have often found the need to annotate relationships, e.g. to add start/stop dates for validity, to indicate data quality, provenance, and so forth. This has encouraged the spread of so called "Property Graphs".

Property Graphs can be defined as graphs formed by vertices and edges where both can have attributes, i.e. sets of name/value pairs, which are often referred to as "properties". Database management systems for Property Graphs currently lack interoperability across different vendors. This is in part due to a lack of formal semantics, unlike RDF which is formally based on description logics, as well the lack of a widely supported standard query language. This is being addressed by ISO work on extensions to SQL, as well as newer work on the GQL query language.

An "ontology" is a term for the data models used to describe graph data. Taxonomies are simple ontologies involving a hierarchy of classes and properties, e.g. as used to classify different species of animals and plants. More complex ontologies involve a richer range of relationships and support simple entailments, e.g. if "tweety" is an instance of the class sparrow, and sparrows are a sub-class of birds, then we can deduce that tweety is a bird. "Knowledge Graphs" is a trending buzzword that refers to graphs used for a combination of data and associated ontologies.

Real-world data is often uncertain, incomplete, and inconsistent. This is a major problem for Data Scientists, as data preparation and massaging can take most of their time rather than mining or modelling data[2]. Moreover, a large majority of data scientists regard data preparation as the least enjoyable part of their work. This creates challenges for combining heterogeneous data sources, something that is especially important to realising the full potential of the IoT.

6.2.2.3 Cognitive AI

People are by and large much better at dealing with the real-world than current computer systems. Perhaps it is time to redress this imbalance by giving computers a human touch! Cognitive AI is AI inspired by advances in the cognitive sciences and holds the potential for doing just that.

Cognitive AI integrates symbolic models (i.e, graphs), statistics; rules and graph algorithms. It moves beyond the limitations of deductive logic by integrating uncertainty and likelihood of different outcomes, e.g. as needed for reasoning about the most likely causes of a faulty machine given the observed symptoms.

[2] https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/#70939cd16f63

Cognitive AI is AI inspired by what we have learned about the brain, and how we can mimic its operation at a functional level on conventional computers, in contrast to Intel's neuromorphic chips which directly simulate spiking neural networks.

Cognitive AI promises a new generation of AI systems with:

- Transparent explanations in terms of statistically grounded knowledge
- Learning from smaller datasets compared to Deep Learning
- Methodologies for combining human expertise with machine learning

W3C 's Cognitive AI Community Group seeks to gather use cases and requirements, to work on a series of demonstrators and outreach for Cognitive AI.

6.2.2.4 The role of data in the development of AI: an EU view

The European Commission launched a consultation on AI on 19th February 2020 inviting citizens and stakeholders to provide their feedback by 14th June 2020.

- "Artificial intelligence (AI) is a strategic technology that offers many benefits for citizens and the economy. It will change our lives by improving healthcare (e.g. making diagnosis more precise, enabling better prevention of diseases), increasing the efficiency of farming, contributing to climate change mitigation and adaptation, improving the efficiency of production systems through predictive maintenance, increasing the security of Europeans, and in many other ways that we can only begin to imagine".

The associated white paper describes the importance of establishing a uniform approach to AI across Europe in order to avoid barriers for the single market. High risk AI applications would be subject to various transparency and other requirements. Low risk AI applications could be subject to a voluntary labelling scheme. The white paper proposes measures to promote an ecosystem of AI excellence, including research funding, skills development, focus on SMEs and partnerships with the private sector. These areas are complementary to the plan set out in the European data strategy.

- "Improving access to and the management of data is fundamental. Without data, the development of AI and other digital applications is not possible. The enormous volume of new data yet to be generated constitutes an opportunity for Europe to position itself at the forefront of the data and AI transformation. Promoting responsible data management practices and compliance of data with the FAIR principles will contribute to build trust and ensure re-usability of data. Equally important is investment in key computing technologies and infrastructures".

6.3 The Emergence of Dataspaces for ecosystems of services

This section reviews the European Data Strategy and its relation to the digital single market across the EU, and the role of data spaces for ecosystems of services [29]. How will emerging technologies facilitate the development of ecosystems of services based upon the IoT? What kinds of standards will be needed?

6.3.1 Metadata: from RDF to Chunks

Today, W3C's Resource Description Framework (RDF) is the dominant standard for expressing metadata. There is a large suite of standards including SPARQL, OWL, SHACL, Turtle, and JSON-LD. There has been plenty of work on ontologies and widespread deployment of the schema.org vocabularies for smart web search results.

The value of RDF has been well proven in many applications in the over two decades since it was first created. However, it is often alleged to be hard for the average developer. This is unfortunate

because it limits wider uptake and prevents RDF from being viewed as a viable choice for many use cases that would otherwise be an excellent fit.

The Easier RDF initiative seeks to build upon community experience with RDF to examine how to make semantic technologies based upon RDF easier to use for the average developer. What aspects or gaps have caused difficulty? How can RDF better support features that users commonly need, and other graph databases offer? How can we make RDF – or a successor – easy enough for *average* developers?

At the same time, businesses are now showing a rapidly growing interest in graph data. Businesses have used relational databases for many years, but it is costly to adapt database schema and applications in response to evolving application needs. Other graph and NoSQL databases have emerged to help meet this need.

Unfortunately, there is a lack of interoperability across existing graph data solutions, motivating interest in open standards for an interchange framework. RDF is an appealing vendor neutral framework for graph data and is well positioned to take on the role of an interchange framework. Although this interest in RDF as a graph interchange framework arose independently from the effort to make RDF easier, and has different goals, there is a natural overlap in motivation, and both efforts can benefit each other.

The guiding principles for Easier RDF are

- The goal is to make RDF (or some RDF-based successor) easy enough for average developers (middle 33%), who are new to RDF, to be consistently successful.
- Solutions may involve anything in the RDF ecosystem: standards, tools, guidance, etc. All options are on the table.
- Backward compatibility is highly desirable, but less important than ease of use.

One candidate for Easier RDF is called “Chunks”, a simple amalgam of RDF and Property Graphs that is under investigation by the W3C Cognitive AI Community Group.

6.3.2 Data spaces and ecosystem of services

In many cases, organisations want to hold onto their own data rather than making it available to others. A common framework for private and shared dataspaces would make it easier for organisations to combine their own services with externally provided services.

Cloud provisioning of storage and compute resources is often cheaper and easier to scale compared to organisations running their own data centres. An example is where emergency field hospitals need to be quickly built and provisioned during a pandemic.

What standards are needed for dataspaces to facilitate ecosystems of services? How can Europe retain the value chain rather than ceding the most profitable parts of it to the Internet giants? These questions remain open and further work is needed.

Data sovereignty is the self determination of individuals and organisations with regards to the use of their data. This aims to restore the balance between the creators of data and the platforms that consume that data. Rather than being asked to sign over ownership or rights over data, data creators can specify data usage policies.

The International Data Space Association (IDSA) has developed a reference architecture around the concept of data sovereignty as a basis for data marketplaces based on European values, i.e. data privacy and security, equal opportunities through a federated design, that ensures data sovereignty for the creator of the data, and facilitates trust among participants.

- *Data providers* grant access to data under specific usage and price models. They are able to control access and usage by the data consumer.

- *Data consumers* can search for data of interest available from data providers. Consumers are bound by the usage policy of the data provider.
- *Connectors* provide standardised connectivity and are responsible for connectivity and usage control. Connectors permit the execution of trusted apps in an isolated identity provider environment.
- *Brokers* manage registrations of data end points exposed by connectors. This allows consumers to look through available data sources and data in terms of their content, structure, quality, actuality and other attributes.
- *Clearing houses* act as intermediaries that provide clearing and settlement services for all financial and data exchange transactions.
- *Identity providers* offer a range of services to create, maintain, manage and validate identity information for all participants in the IDS architecture.
- *App stores* provide data apps, i.e. applications that can be used for tasks like data transformation, aggregation and analytics. Apps may be certified by IDS approved certification bodies. App stores are provided by IDS members.
- *Vocabulary providers* manage and offer vocabularies (ontologies, reference data models, metadata elements) which can be used to annotate and describe data sets.

IDS is seeking to support a wide range of data: mobility data, technical drawings, sensor data, financial data, material features, medical data, quality data, origin data, planning data, geographical data and many others.

Further work is needed to explore opportunities for data spaces based upon Cognitive AI and Cognitive databases. One area of interest is to enable fine grained control of data sovereignty, i.e. on sub-graphs within a graph database, rather than on the database as whole. For example, considering how to record derivation chains (data lineage) for operations that transform data items.

6.3.3 The relationship between data spaces and the IoT

The Internet of Things provides network access to physical devices with sensors and actuators. For sensors, data can flow through successive stages of interpretation, and exposed to ecosystems of services via data spaces. Applications may be interested in the "current" value of some measured value, the history of sensor readings, and the means to respond to specific events, e.g. when a measured value crosses a given threshold. Data spaces need to cater for these different needs by providing a suitable application API.

In some situations, applications will need to adjust sensor parameters. Applications will also want to be able to invoke operations on actuators, either as individual operations, or using a stream of updates, e.g., to control the joints of a robotic arm. Data spaces thus need to enable such operations to be sent back to the IoT devices.

This can be modelled in terms of the Web of Things which provides an abstraction layer for digital twins. Each thing is given a unique identifier for use in rich descriptions of the object model exposed to applications, in terms of properties, actions and events, as well as the semantic context, e.g. the kind of thing, its function, location and other metadata.

In summary, data spaces can be used to expose digital twins as a basis for remote interaction with IoT devices hosting sensors and actuators.

6.4 Privacy-based Business Models

This section starts with a look at how behavioural tracking of users' online activities have supported business models based upon advertising or improved sales. This is often perceived as an unwarranted intrusion into personal privacy. Moreover, end-users are often deluged by

advertising for something they have recently purchased and have no interest in a further such purchase.

This suggests a different approach that effectively turns privacy on its head in the sense of providing servers with rich personal details on demand to obtain higher value services, where some of these details can be based upon information collected from IoT devices, e.g. the current location of users. This approach can be considered as pull-based services as contrasted with the push-based services associated with advertising.

6.4.1 Behavioural Tracking

The Web today is dominated by advertising-based business models that depend upon behavioural tracking of the end users. One of the best known tools is Google Analytics which is used by organisations to monitor visitors to their website. Behavioural tracking can be used to spur on customers towards a purchase, e.g. by providing recommendations based upon purchases by other customers, by offering a trial or a discount, by offering a downloadable white paper, and by sending potential customers further marketing materials to their email address.

Many web sites include advertising provided by an advertising agency that is dynamically chosen based upon profiling the end-user. In some cases, this process is determined using a real-time bidding market. The web page directs the browser to contact the advertising server with a request for advertisements along with information about the user, such as cookies and the page being viewed. The advertising server then contacts a supply-side platform which in turn contacts a data management platform responsible for collecting information about the particular user.

This data includes first-party data from customer relationship management platforms, second-party data using statistics related to cookies on advertising servers, social and analytics platforms, and third-party data obtained from external providers that purchase data from businesses. The supply side platform is then able to pass this information to an advertising exchange that supplies the call for bids to demand-side platforms that are expected to rapidly decide how much to offer for the advertising space. For more details, see the Interactive Advertising Bureau.

Whilst many end-users are content to be tracked across the Web, many others are uncomfortable with seeing advertisements relating to their online behaviour. This has had several consequences. Web browsers now commonly enable users to block third party cookies (information set by a third-party server to track users across visits). Companies have responded by finding other ways to track users based upon fingerprinting their browser. W3C Working Groups now pay careful attention to how each browser API contributes to finger printing.

Privacy concerns stimulated the development of ad blocking software which has been widely adopted, e.g. a survey in 2016 reported that 72 million Americans were using ad blockers. Initially, ad blockers were implemented as browser extensions, but more recently this has been discouraged by browser vendors in favour of pre-determined filters built into the web browser. Other approaches include the use of a web proxy that filters out content before it can reach the browser. This can further be applied to blocking content that may be offensive, inappropriate or malicious. Proxy based solutions are unable to filter traffic routed by transport layer security. Another related approach is to filter by the server's domain name (DNS filtering).

W3C launched the Tracking Preference Expression (DNT) Working Group in an attempt to standardise an HTTP header field designed to allow users to opt out of tracking. Whilst this is now widely supported by web browsers, it failed to be widely supported by industry, and the Working Group was closed in January 2019.

The European Commission introduced the requirement for websites to get end-user approval for online tracking under the GDPR regulations. Users must be able to opt-out of all except strictly necessary cookies and still be able to use the website. GDPR also enshrines the right to be

forgotten, i.e. for end-users to request the erasure of personal data without undue delay, under certain specific circumstances. In addition, end-users must be able to access their personal information and request corrections to any errors.

6.4.2 Pull-based services

End-users sometimes perceive websites to be a little "spooky" when they see page content based upon their recent online behaviour on other websites. Furthermore, there are diminishing returns for behavioural tracking as it is often backwards facing – describing the user's past behaviour rather than what the user is seeking to do next. This suggests that it is now timely to consider other approaches in which users volunteer rich personal details in return for the high value services they are looking for.

One example is for travel, where people are trying to arrange a holiday or business trip. They are likely to want to know much more about the destination such as how to get to the departure airport and how to get from the destination airport to their hotel, whether it is worth getting a travel card or what kinds of coins they are likely to need for trips across town by bus and metro. Where to eat near their hotel, what to see and so forth. Having to interact with many different websites to gather this information is both time consuming and frustrating.

Another example is where people are seeking medical advice specific to their circumstances, which involves providing access to that person's medical history, as well as to data collected from IoT devices such as fitness bands, smart weighing scales and so forth, including their food purchases and lifestyle behaviours. This is likely to be of increasing importance as people are living longer, and we want to enable people to live better as they get older and frailer, and set against the context of the desire to keep costs from spiralling up, and limits on the number of caregivers etc.

A user-centric view of privacy can be supported by a trusted third party that manages personal information on behalf of the user, and which provides appropriate information on demand to the services requested by the user, and subject to terms and conditions that safeguard data sovereignty for the user's data.

Such data can include data from smart home sensors, including wearables, and potentially devices outside of the home that can gather data from the environment around them, e.g. video cameras. In such cases, the users are granting temporary permission to use data that relates to them. This raises technical and regulatory questions around how public IoT devices can identify someone as a precondition for enabling access to third parties on behalf of that person, without unauthorised access to information on other people.

Many users will prefer to delegate decisions over the details on what personal information to disclose and under which circumstances. For such users, the privacy manager can make an informed decision that reflects the user's risk profile and values as determined by an analysis of the user's stated preferences, the user's past behaviours, and those of people like him or her.

Users may also be prepared to give service providers the means to operate actuators, e.g. to control home lighting and home entertainment systems, as is the case for today's smart speakers e.g. Amazon Echo.

This points to the need for a framework for certifying services as trustworthy, along with machine interpretable descriptions of the personal data they need, and a means to negotiate which such data is made available, and the details for the agreed terms and conditions. There is a need for auditable records of transactions that can be used in disputes.

A further opportunity is for the provision of a bidding system in which rival services compete to fulfil the user's request, where the privacy manager ranks the bids according to a variety of dimensions including price, popularity and features offered.

W3C is now at an early stage of planning a Workshop on privacy-based business models to bring together representatives from different communities, e.g. healthcare and web marketing, to share ideas and to identify specific opportunities for standardisation.

7 SUMMARY AND CONCLUSIONS

7.1 Lessons learned

The IoT spans many domains, technologies and communities, and the level of fragmentation we see today is therefore unsurprising.

The experience gained with the Large-Scale Pilots shows that convergence is possible, and that a common reference model can help with the dialogue across communities and across domains.

The LSPs and Coordination and Support Actions (CSAs) along with the AIOTI have boosted liaisons between companies, research institutions, industry alliances, standards development organisations and regulatory bodies.

The IoT has a prominent role to play in digital transformation along with work on Data, Artificial Intelligence and Machine Learning, Distributed Ledgers, and ecosystems of services based upon Data Spaces.

For communication engineers, there are challenges around the role of 5G and the Next Generation Internet technologies. For application developers, the focus is on scalable approaches to data and metadata, and how we can build secure, trusted value chains for ecosystems of services that unlock the full potential of the IoT.

7.2 A summary of LSP contributions

The five LSPs (of the first generation) of LSPs has made a great number of contributions to standardisation in a variety of SDOs and SSOs. These contributions were, in most cases, supported by a Standardisation Plan defined upfront at the beginning of each LSP project and further elaborated with the progress of the work.

A large part of these contributions has been going to IoT standardisation and some examples are listed below:

- The collaborative development by LSPs of a 3D Reference Architecture model.
- The Minimum Interoperability Points (MIMs) specified by ACTIVAGE, IoT 2020 and SynchroniCity.
- The MONICA requirements for a new standard for time-critical data links for IoT sensors.
- The LSPs contributions to SAREF (Smart Appliances REFerence ontology).
- AUTOPILOT contributions to oneM2M
- The contributions of SynchroniCity to the ITU Study Group 20 on IoT and Smart Cities.

In addition, the LSPs have developed methods and tools in support of the efficient implementation and deployment of the LSP Use Cases that may possibly become more standardised elements to be used by the second generation of the LSPs, such as:

- The IoT Catalogue that collects components that can be used and reused in implementation projects.
- The methodology for launching Open Calls and following their implementation.

7.3 Future Work for the IoT Standardisation Community

Innovation in the IoT field is continuing at high pace and priorities have somehow shifted, in particular due to the central role of data in IoT systems and the growing role of Artificial Intelligence in dealing with all issues related to the management and processing of data.

Some topics have been clearly identified (see deliverable D06.11 [10]) as priorities by several projects of the second generation of LSPs.

A short list of these topics is the following:

- Supporting the efficient adoption of new technologies such as Distributed Ledger Technologies or Artificial Intelligence, across all parts of the IoT systems from physical to business layers, in support of cross-cutting functions such as security, privacy or safety.
- Developing solutions (including the necessary infrastructure) for secure data management in support Open Access to data and the creation of generic or sector-specific data spaces.
- Addressing the challenges of industrial adoption for semantic interoperability together with improving the efficiency of organisational interoperability solutions.
- Strengthening the provision of privacy and security for resilient services.
- Taking full benefits of communications scalability, reliability, latency (e.g., 5G).
- Boosting the efficiency of edge solutions (data privacy, federation, AI, etc.).
- Reducing fragmentation around auto-configuration and discovery in ways that preserve security and privacy, and abstract away from heterogeneity of underlying systems.
- Proposing robust solutions for privacy-preserving federated machine learning.
- Fostering the emergence of the Intelligent autonomous IoT.
- Defining the Sentient Web (web of digital twins, cognitive AI and open marketplaces).
- Fostering the development of common visions (e.g., through White Papers) across the largest possible number of relevant actors in the field of standardisation.
- Promoting recommendations for collaborations across PPPs, SDOs and other alliances.

It is expected – and an achievable target – that the next generation of LSPs may contribute to the progress in IoT standardisation, even on a larger scale than the first generation has. To make this most effective, there is a need, across the new LSPs, for collaboration, information exchange, as well as identification and promotion of common approaches and best practices.

8 REFERENCES

CREATE-IoT WP06 Deliverables:

[1] "Strategy and coordination plan for IoT interoperability and standard approaches", Deliverable D06.01, 2017.
[2] "Recommendations for commonalities and interoperability profiles of IoT platforms", Deliverable D06.02, 2018.
[3] "Assessment of convergence and interoperability in LSP platforms", Deliverable D06.03, 2020.
[4] "IoT pre-normative activities", Deliverable D06.04, 2017.
[5] "Initial report on IoT standardisation activities", Deliverable D06.05, 2018.
[6] "Strategy and coordination plan for IoT interoperability and pre-normative and standardisation activities", Deliverable D06.07, 2018.
[7] "Interoperability Framework Workshop", Deliverable D06.08, 2018.
[8] "Workshop on LSPs use cases: integration and standardisation alignment", Deliverable D06.09, 2019.
[9] "Workshop on IoT standardisation activities", Deliverable D06.10, 2019.
[10] "Workshop on common IoT standardisation framework", Deliverable D06.11, 2020.

Other CREATE-IoT Deliverables:

[11] "EU research and innovation activities overall plan", Deliverable D04.03, 2017.
[12] "IoT Policy Framework", Deliverable D05.01, 2017.

LSP References and Deliverables:

[13] ACTIVAGE (ACTivating InnoVative IoT smart living environments for AGEing well); https://european-iot-pilots.eu/project/activage/
[14] "ACTIVAGE IoT Ecosystem for Smart Living Environments", ACTIVAGE Brochure
[15] AUTOPILOT (AUTOmated driving Progressed by Internet Of Things); https://european-iot-pilots.eu/project/autopilot/
[16] "Standards and conformance of IoT in AD", AUTOPILOT Deliverable D5.8
[17] IoF2020 (Internet of Food and Farm 2020); https://european-iot-pilots.eu/project/iof2020/
[18] "Hosting Environment and IoF2020 Lab", IoF2020 Deliverable D3.8
[19] MONICA (Management Of Networked IoT Wearables); https://european-iot-pilots.eu/project/monica/
[20] SynchroniCity (Delivering an IoT enabled Digital Single Market for Europe and Beyond); https://european-iot-pilots.eu/project/synchronicity/

Other References:

[21] "Advancing IoT Platforms Interoperability", River Publishers, Gistrup, 2018, 978-87-7022-005-7 (ebook), IoT European Platforms Initiative (IoT-EPI) White Paper, online at: https://iot-epi.eu/wp-content/uploads/2018/07/Advancing-IoT-Platform-Interoperability-2018-IoT-EPI.pdf
[22] oneM2M TR-0026 "Vehicular Domain Enablement", see: http://onem2m.org/technical/published- drafts/release-4
[23] AIOTI report "IoT relation and impact on 5G", see: https://aioti.eu/aioti-report-on-iot-relation-and- impact-on-5g/
[24] Vermesan, O. and Bacquet, J. (Editors). *Next Generation Internet of Things - Distributed Intelligence at the Edge and Human Machine-to-Machine Cooperation*. ISBN: 978-87-7022-008-8 (Hardback), 978-87-7002-007-1 (Ebook). River Publishers, 2018.

[25] Alliance for Internet of Things Innovation (AIOTI). Identifiers in Internet of Things (IoT), version 1.0. AIOTI WG3 - IoT Standardisation, February 2018. Online at: https://euagenda.eu/upload/publications/identifiers-in-internet-of-things-iot.pdf

[26] The Industrial Internet of Things: Managing and Assessing Trustworthiness for IIoT in Practice. White Paper, version 1.0. Industrial Internet Consortium July 2019.

[27] International Organization for Standardization (ISO). Online at: https://www.iso.org/home.html

[28] Handbook to the IoT Large-Scale Pilots Programme. Online at: https://european-iot-pilots.eu/resources/iot-european-large-scale-wiki/

[29] Roy Fielding's Doctoral Dissertation on Architectural Styles and the Design or Network-based Software Architectures, see: https://www.ics.uci.edu/~fielding/pubs/dissertation/top.htm

[30] The European Digital Strategy, https://ec.europa.eu/digital-single-market/en/content/european-digital-strategy