ITEA3 18008
TIoCPS project

A state of the art related to scope of TIoCPS project

TIoCPS
ITEA 3 18008

Edited by Juhani Latvakoski/VTT

Including contributions from:
Bittium/Jussi Polet
Sirris/Annanda Rath
Alpata/Murat Saglam
ISEP/Luis Gomez
VTT/Juhani Latvakoski
Elvak/Markus Kiviniemi

TABLE OF CONTENTS

TABLE OF FIGURES

Abstract

The motivation for TIoCPS project arises from the grand challenge facing cyber-physical systems (CPS). The lack of digital trust prevents the establishment of information sharing around cyber-physical systems (CPS), and thus establishment of the data economy around CPS. The objective of the project has been to technically enable trustworthy and smart communities for CPS systems (TIoCPS concept) for solving the referred grand challenge in the context of selected industrial use cases dealing with energy, mobility and user/owner CPS systems. The targeted TIoCPS concept is envisioned to result in the form of an interoperable CPS based real-time ecosystem, especially around the focused use case solutions, which is boosting the businesses of the respective industries, thus fostering a more smart, sustainable and interoperable future society. The aim of this deliverable is to represent a view to the state of the art related to TIoCPS project scope.

List of Abbreviations

Abbreviation	Explanation
A	agreements
AA	authorization and authentication
ABAC	attributes-based access control
API	application programming interface
BACnet	building automation and control networks
BLE	Bluetooth low energy
CCTV	closed-circuit television
CPS	cyber physical system
DALI	digital addressable lighting interface
DL	distributed (blockchain) ledger
DS	discovery services
EIF	European interoperability framework
ENS	exposure notification systems
ESA	external service access
EU-GDPR	European Union general data protection regulation
GATT	generic attribute profile
GDPR	general data protection regulation
HSM	hardware security module
HTTP	hypertext transfer protocol
IAM	identity and access management
IDS	international data spaces
IDSA	international data spaces association
IoT	internet of things
IPFS	interplanetary file system
JWT	json web token
MDS	mobility data space
Modbus/RTU	Modbus / remote terminal unit
Mosbus/TCP	Modbus / transmission control protocol
MQTT	message queue telemetry transport

OPC/UA	open platform communications unified architecture
OTP	one-time password
PEP	policy enforcement point
PKI	public key infrastructure
PLC	programmable logic controller
RBAC	role-base access control model
REST	representational state transfer
SC	service components
SHSD	secure handling of sensitive data
SSI	self-sovereign identity
TCP	transmission control protocol
TCP/IP	transmission control protocol / internet protocol
TIoCPS	trustworthy and smart communities of cyber-physical systems
UDP	user datagram protocol
URL	uniform resource locator
VC	verifiable claim
XACML	extensible access control markup language

1 Introduction

The motivation for TIoCPS project arises from the grand challenge facing cyber-physical systems (CPS). The lack of digital trust prevents the establishment of information sharing around cyber-physical systems (CPS), and thus establishment of the data economy around CPS. The objective of the project has been to technically enable trustworthy and smart communities for CPS systems (TIoCPS concept) for solving the referred grand challenge in the context of selected industrial use cases dealing with energy, mobility and user/owner CPS systems. The targeted TIoCPS concept is envisioned to result in the form of an interoperable CPS based real-time ecosystem, especially around the focused use case solutions, which is boosting the businesses of the respective industries, thus fostering a more smart, sustainable and interoperable future society.

The aim of this deliverable is to represent a view to the state of the art related to TIoCPS project scope. Because the scope is quite a large, it has been important to divide the content in a way or other into some understandable form. The selected division in this deliverable is based on the original work plan structure of the project. According to it, the rest of this deliverable is structured so that the chapter 2 opens a view to the overall architectures of Cyber-Physical Systems. Chapter 3 discusses about the state of the art and practises around information level technologies and applications. Chapter 4 provides a view to the communications technologies and applications. Chapter 5 discusses about the CPS devices, trust and security. Finally, some concluding remarks are explained in the chapter 6.

2 Review of the Cyber-Physical Systems architectures

The aim of this chapter is to open a view related to the state of the art and practises from the architectural perspectives of the Cyber-Physical Systems.

2.1 Technologies

There are several specifications, industrial forums and even standards targeting industrially relevant architectures of cyber-physical systems. Separate specifications have been developed for each sector/vertical domain, such as e.g. home/buildings, manufacturing/industry automation, vehicular/transportation, healthcare, energy, cities, wearables etc [41]. However, there are also initiatives aiming at cross-domain, horizontal type of IoT platforms [1, 2, 3, 4, 5, 6, 7, 8]. In addition, several standardization bodies like IEEE, OMG, W3C, OpenFog, AllSeen alliance, NGMN and ISO/IEC JTC1 WG41 have been working in the area. There are also other related actions such as e.g. securing IoT products with blockchain [9], and comparisons and related studies of IoT Platforms. For example, Guth *et al.* compare OpenMTC, Fiware, SiteWhere, AWSIoT and provide IoT specification with IoT integration middleware, Gateway and Devices as basic building blocks [10]. Burg *et al.* focuses to review wireless communications and security technologies for cyber-physical systems and conclude that security is an essential challenge for wireless cyber-physical systems operating in horizontal way across multiple domains [11].

Essential solutions in the referred horizontal IoT specifications are related to the edge system and IoT platform. The edge system usually comprises identifiable physical entities, which can be connected to IoT infrastructures and platforms either directly or via some sort of gateway [12, 13, 14, 15]. An IoT platform is typically an integrated physical/virtual entity system capable of controlling, monitoring, information processing and application execution. There are also typically different kinds of tiers defined according to the accessibility of the entities, platform and enterprise systems. The information models are used to define the properties of IoT information content. In addition, several studies have investigated how virtualization capabilities of IoT systems can be deployed at the edges of the network.

There are a number of challenges that are not properly addressed. One challenge concerns the collaboration between consumer and industrial endpoints, which is not supported properly by any of the existing specifications. There are also fundamental requirements towards crosscutting solutions in the areas of security, safety, interoperability, composability, data management, analytics, resilience, composability, virtualization, and regulation. It is obvious that creating horizontal solutions have essential benefits, because they can be deployed in multiple domain scenarios, and minimizes the need for application domain specific solutions [16, 41]. However, this is especially challenging when speaking about the information and service level, which easily mix the domain and potentially horizontal generic services when handling services related to information streams. In addition, the dynamic changes in IoT systems, such as continual adding of physical entities involving OEMs and related SPs, heterogeneous sensors and actuators and other devices, have proved to be challenging especially from a security point of view. The Industrial IoT devices are often used in physically protected and isolated environments; however, today there is the need to enhance the operation also with many other devices (e.g. for energy flexibility). State-of-the-art IoT specifications lack proper solutions for solving this challenge. Proper identification, authentication and authorization capabilities seem to be missing for dynamic IoT environments, which prevents establishment of trust relationships. Uncertainty exists in information ownership and validity, and to remote management of the physical assets. Therefore, the lack of digital trust in communications between physical cyber-physical resources owned by different stakeholders have proved to be one of the most essential challenges.

The rest of this section opens shortly some potential state of the art architectures and technologies that are estimated to provide some partial solutions for the challenges described and are estimated to be relevant for the TIoCPS project scope when writing this report.

2.1.1 GAIA-X

GAIA-X is a European initiative to develop a digital governance for transparent, controllable, portable and interoperable data exchange for existing cloud/edge technology stacks based on European values [17]. GAIA-X recognizes the importance of trust establishment as a facilitator to connect various services in unprecedented ways. Their approach is to avoid monolithic, centralized authorization and rather approach the problem via the use of decentralized, federated identifiers and services. Those, in turn, are enabled by pervasive use of self-descriptions that consist of claims whereby the assets describe their relevant features in a way that can be (e.g., cryptographically) verified and subjected to being catalogued and queried for discovery purposes.

GAIA-X does not provide centralized data storage, either. Instead, each company decides for itself where its data is stored, as well as who may process it and for what purpose [18]. GAIA-X implements mechanisms by which data and service providers can authorize and control the usage of the data. However, GAIA-X does not implement transaction-level monitoring; it is entirely the responsibility of the data provider to ensure that all data use is within the agreed policies. Thus, GAIA-X does not guarantee, for example, GDPR compliance of the services, other than controlling data source access and usage policy.

However, GAIA-X is still in its early phases and it is not fully clear how it will be suited for temporary, ad-hoc data exchange. The threshold for adopting GAIA-X has been kept low so that it is equally suited for startups and SMEs as well as larger corporations. Still, the process for onboarding providers, services and data could be laborious for short-term purposes – at the minimum it is required that all participants and assets are reliably identified, described using self-descriptions and policies, determined compliant and approved by GAIA-X, and deployed on GAIA-X compatible nodes and services. How streamlined this process can be remains to be seen; for participants and assets that are already in GAIA-X, the burden should not be overwhelming.

2.1.2 MyData

The MyData [19] approach leans towards decentralized operation, which suggests the use of self-sovereign identifiers (SSI) with technology support provided by operators to help users remain in control of their data. Such support could include, e.g., tailored user interfaces for ease of use, and ready-made privacy and data sharing profiles for various groups of users. Community-based profiles, and the related “collective” business model, reflect the TIoCPS concept rather well.

The MyData data sharing model with data primarily flowing directly between services is also relevant to TIoCPS. MyData approaches the related authorization and permissions via explicit permission / consent documents that are kept separate from the data flow; it is the responsibility of each data operator to verify using the consent documents that each requested operation is allowed.

While MyData is largely a concept rather than ready implementation, it discusses many topics that are central to TIoCPS. MyData does not mandate or provide any specific technical solutions for its implementation. However, a reference architecture is available, as well as some examples of how the functionalities could be implemented. Furthermore, 41 companies have already been awarded the status of “MyData operator” as of April 2023. The MyData model enables different scenarios for organizing personal data infrastructures. However, there is strong support in the MyData community towards a model

where there are multiple competing operators that provide privacy-aware services to individuals in a globally interoperable fashion, much like current telecom operators, energy providers, or banks. This would be in line with the TIoCPS concept of different companies offering their specific functionalities into the common ecosystem.

2.1.3 International Data Spaces

IDSA – The International Data Spaces Association – aims at the development of a global standard for international data spaces (IDS) and their interfaces, as well as to foster the related technologies and business models that will drive the data economy of the future across industries.
In April 2023, the IDSA consists of approximately 150 member organizations.
The number of organizational and company members in the IDSA is constantly increasing. Their Reference Architecture Model is fairly comprehensive, which is also attested by the formalization of the IDS Connector model into the DIN SPEC 27070 standard, which is also expected to be a central element of the GAIA-X architecture [20]. The connector as a component for joining heterogeneous partner data sources to a common ecosystem is central in TIoCPS as well. The IDS connector model not only enables data transactions directly between the data provided and its user (via Connectors) but interestingly also allows running data user's code at the data source Connector.
One notable development is the Dataspace Connector, an open-source project that is being developed in partnership with several research institutes and businesses. Its architecture makes it possible to alter the existing implementation to meet domain-specific requirements. Existing applications can be readily extended by IDS connector functions and integrated into an IDS data ecosystem using the Dataspace Connector. Furthermore, the Dataspace Connector can be used as a foundation for developing custom software that connects to an IDS data ecosystem [21].
The Dataspace Connector provides a REST API for managing datasets as IDS resources based on their metadata. External data sources can also be connected to the Dataspace Connector via REST endpoints, allowing the Dataspace Connector to function as a middleman between the IDS data ecosystem and the data source itself. The Dataspace Connector can act as both a data provider and a data consumer at the same time, allowing it to both provide and request data in a data ecosystem. Various usage control rules are implemented and enforced by the Dataspace Connector. This enables usage control rules to be assigned to data in the IDS data ecosystem, ensuring data sovereignty throughout the data lifecycle. Furthermore, the inclusion of an identity provider in the IDS context, such as a DAPS, aids identity management.
IDS puts a lot of effort in ensuring the trustworthiness of participants in data exchange transactions, and in providing the necessary access control, usage control and monitoring services to make sure that contracts are respected. All this is largely achieved via rigorous screening of potential participants and excessive certification via a centralized certification authority. Certificates and the necessary processes to ensure that they are only granted to participants that express due rigor in their security processes are extended to all infrastructure components and their operators. Respecting individual users' rights to control the use of their data seems to be well covered through explicit data usage permissions and enforcement of their use, and comprehensive provenance tracking.
The IDS ecosystem model mandates that trust in an arbitrary transaction must be extensively built a priori – to the participants, their data connectors, any data operators used, usage rules, etc. While this might not be an overwhelming task for partners and datasets that are already onboard the IDS, it is not clear how arduous the process would be when the usage should be changed and agreed on a more or less ad-hoc basis. IDS is conservative in its approach in that it is based on centralized certification and formal processes, which leads to a high level of trust but a heavy onboarding process. One attestation of the

centralized approach is that blockchain technologies are only cursorily mentioned in the IDS-RAM; they are expected to play a role in maintaining shared data assets in an IDS environment, such as verification of datasets via blockchain-stored hashes. The Clearing House and Broker services are also expected to benefit from blockchain technologies, but the core IDS approach relies on centralized identification, authentication, and authorization.

2.1.4 Energy Web

Energy Web [22] put data sharing to the center of future energy solutions where electricity is increasingly generated by fragmented, often unpredictably available renewable production and consumer-controlled, distributed energy production and consumption control resources. This is fully in line with the TIoCPS concept and particularly with the energy-related use cases. The Energy Web solutions also address central questions in TIoCPS, such as how authentication, user/resource identification and data sharing practices and limitations can be implemented in a way that the end consumers have the final say on sharing their data, and personal data is shared to the minimum amount possible. Exposing just enough information to establish the necessary amount of trust for the intended transaction is key.

With their EW-DOS, Energy Web are building a platform that enables loosely coupled applications, services, and partners to work together for compound goals. There is no central authenticator, as the solution is fundamentally based on decentralized, blockchain-anchored identities, decentralized identifiers, verifiable claims, and smart contracts. While the platform is primarily aimed at the energy sector, it is based on general use, open-source components (e.g., Ethereum, W3C DID).

Decentralized technologies by design cover some of the requirements listed for TIoCPS. User control over their data is the core principle of self-sovereign identities. With DIDs and verifiable claims, trust establishment is fine-grained and based on selective exposure of personal data; for example, if a service is only interested that a user is of legal age, a verifiable claim could provide just that and nothing more, not even the user's age or date of birth. In many TIoCPS scenarios, claims-based user or device attribute sharing could be one way to establish minimum required trust.

Another benefit of self-sovereign identities is that service providers do not necessarily have to collect personal data, as it is controlled by the user. It would greatly ease GDPR compliance, if the service provider does not have any information that can be traced back to an individual person. Similarly, SSIs make data sharing to third parties visible to the user by design, which is another requirement discussed in TIoCPS.

As identity and credentials management explicitly become the responsibility of the user, they can become burdening. For this purpose, the Energy Web are developing the EW Switchboard [23] that is an open-source application [24] for identity and access management (IAM). It focuses on the energy market, but according to EW can be used in any sector, whether the solutions are based on blockchain or not. The EW Switchboard leverages SSIs, DIDs and VCs for authentication and authorization (for example, role-based access management), and user activity logging.

From TIoCPS viewpoint, the Application Registry feature of the EW-DOS is particularly interesting. An Application Registry consists of identities (users, assets) that, via verifiable claims, fulfil some predefined criteria such as geographic region, and implement administrative features specific to those criteria. Such a registry could be seen to implement, at least in part, the TIoCPS concept of a community. Plus, as the members have already been verified to comply with community rules, it would automatically meet the need for prior agreement negotiation, which is especially needed in some TIoCPS energy use cases where the need for demand-response actions may arise very quickly.

All in all, there do not seem to be active plans for using the EW-DOS or the EW Chain for other sectors than energy. However, the technologies are based on open-source components and open standards, so adapting parts of the Energy Web approach for more generic use cases should be possible.

2.1.5 Mobility Data Space (MDS): A Secure Data Space for the Utilization of Mobility Data

Mobility Data Space (MDS) [25, 26] is an open data space platform being created which offers access to traffic data and sensitive mobility data beyond their secure exchange. This platform links existing data platforms to each other and allows them to exchange and share data securely and seamlessly. This platform is currently running at regional level in Germany. However, in the future, it will thus be possible to provide comprehensive mobility data on a national level.

MDS adopts a decentralized system architecture developed by the International Data Spaces Association [27]. MDS offers an ecosystem in which data providers can specify and control the conditions under which their data can be used by third parties. This approach creates data sovereignty as well as trust, and data users can be sure about data origin and quality.

MDS's Architecture is established across the networked connectors (interfaces for different platforms to connect), it is not a centralized platform but rather an expandable network of decentralized players. Prior to being transferred to the target connector, the data to be provided is extended by a set of rules, the so-called "usage policy". It remains in the target connector and is secure against direct access by the data user. If data users want to work with the data, they must access it within the connector via so-called data apps which are capable of integrating further data. There is also a usage control within the connector. This ensures compliance with the rules specified by the data app, with the result that only aggregated results will leave the connector.

The Mobility Data Space [28] is the data sharing community for those who is looking to build future mobility services. Its goal is to facilitate competition around innovative, environmentally sustainable, and user-friendly mobility concepts by offering all players/users equal and transparent access to relevant data. All users, within data space, have unique opportunities to benefit from the added-value potential of their data. The Mobility Data Space promotes the development of forward-looking mobility services - based on high standards of data protection and data security "Made in Europe". This EU data space covers wide range of mobility use cases in different areas: from vehicle manufacturers to ride-share services, from public transport operators to navigation software companies, from research institutes to bike-sharing companies.

2.1.6 Fiware

FIWARE [29] is a platform aims to manage context data in a generalized set of standards with the use of its APIs to implement in smart solutions. Context data are the virtual representations of the real-world objects, people, and relationships between them. FIWARE components are open-source, and the middleware for the platform is the Orion Context Broker. Orion Context Broker provides an API for managing context data that is called NGSIv2 API.

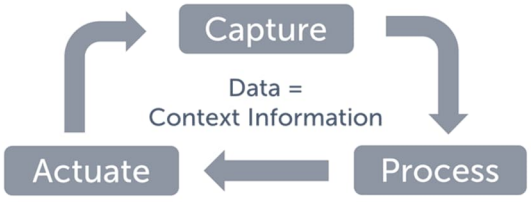

Figure 1. Context Data Flow (FIWARE, 2020).

FIWAREs other components support the context broker in terms of:

- supplying context data from various sources (IoT, social networks, robots),
- managing context data,

- processing, analyzing and visualization of context data,
- accomplishing complex event processes,
- authorization, access control and monetization.

Following services can be used for receiving-sending data, recording, visualizing and analyzing data. Additional services may be added later. The services and operating systems mentioned below are all open-source and community driven which requires no purchase.

1. Orion Context Broker (FIWARE), middleware for holding the latest state of the virtual entities and sending updates to other services, databases with subscriptions.
2. MongoDB, no-sql database that will be used by Orion and Draco. Orion will store virtual entities and subscriptions. Draco will store past data of these virtual entities.
3. Draco, an alternative data persistence mechanism for managing the history of virtual entities.
4. Mosquitto, message broker that will implement the MQTT protocol for the electrical motor.
5. ROS, robotics middleware that will be used in robot.
6. FIROS, tool for translating ROS messages into NGSI to publish them in Orion.
7. IoT-Agent-Ultralight, an IoT Agent that will translate MQTT messages into NGSI to publish them in Orion.
8. User Interface: Interface for making http requests on Orion Context Broker

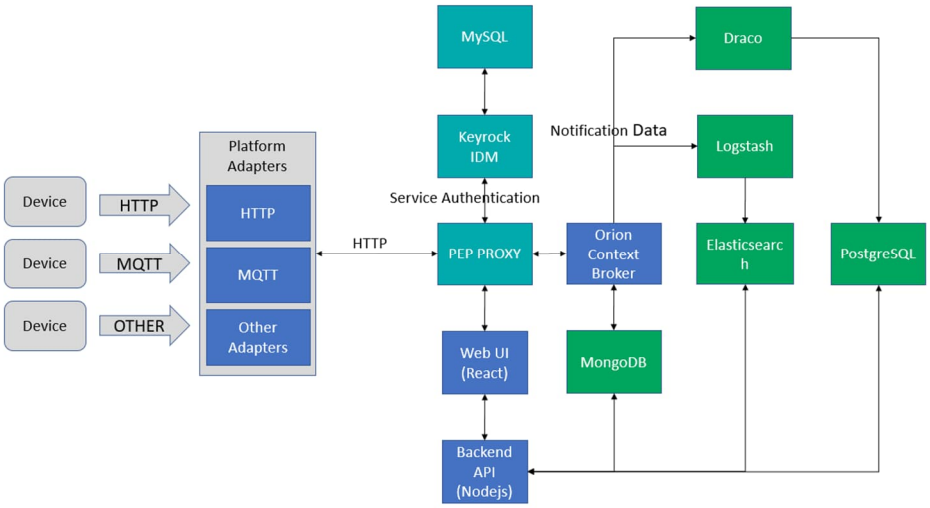

Figure 2 General data flow Fiware infrastructure.

2.1.7 ArrowHead

Arrowhead [30, 31] is a framework composed of local clouds, devices, systems, and services. The main goal of the arrowhead framework is to achieve interoperability between heterogeneous system using existing protocols for handling legacy systems. Arrowhead has been applied in numerous IoT automation scenarios [32], such as the efficient deployment of a large number of IoT sensors, programmable logic controller (PLC) device monitoring, replacement devices, energy optimization, and maintenance.

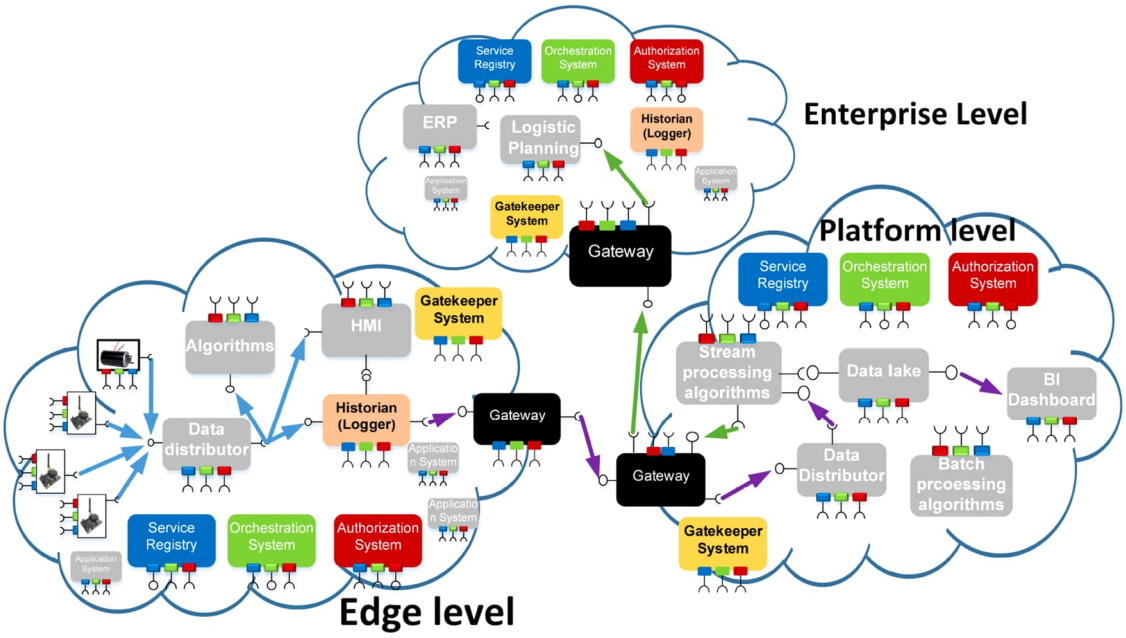

Figure 3. Example of arrowhead structure.

2.1.8 FlatTurtle

Flat turtle [33] is a small (commercial) API that allows to visualize relevant information around a building. It connects data spaces but without the service layers. The connectors are links to internet web sites. It does not connect to CPS edge devices but uses the Wifi of the building visitors to collect and provide data.

2.1.9 European Interoperability Framework

The European Interoperability Framework (EIF) gives specific guidance on how to set up interoperable digital public services. The purpose is to setup a single digital market. The EIF focusses on public administrations. Its guidelines can be generalized but are more related to administrative data then CPS. Conceptual model of EIF is depicted in Figure 4.

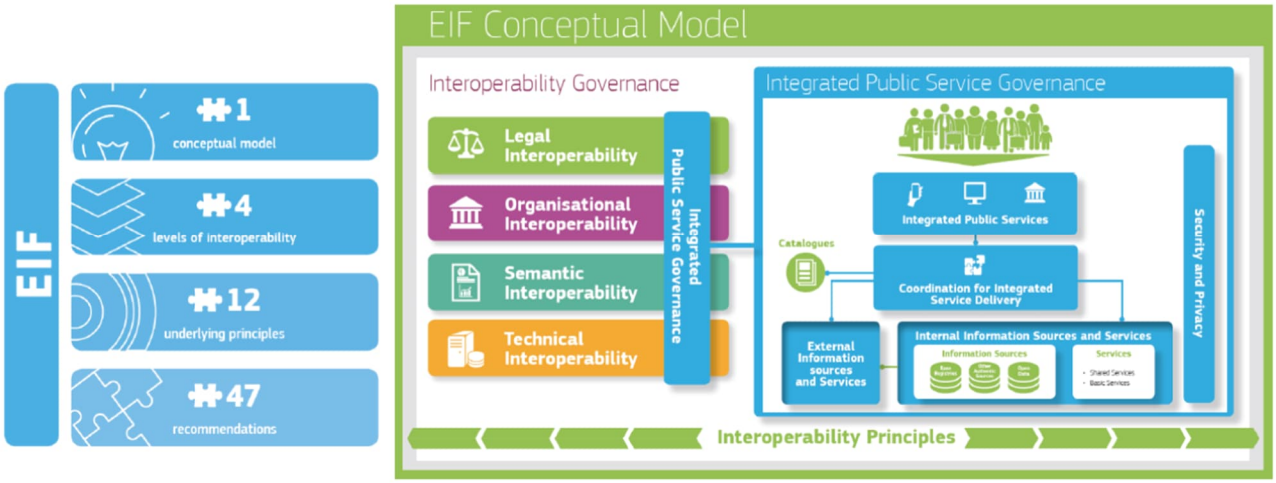

Figure 4 EIF conceptual model.

2.1.10 Wikidata

Wikidata [34] is a free and open knowledge base that can be read and edited by both humans and machines. Wikidata is a collaboratively edited multilingual knowledge graph hosted by the Wikimedia Foundation. It is a common source of open data that Wikimedia projects such as Wikipedia can use under the CC0 public domain license. Wikidata is a wiki powered by the software MediaWiki. It is also powered by the set of knowledge graph MediaWiki extensions known as Wikibase.

Wikidata is a document-oriented database, focused on items, which represent any kind of topic, concept, or object. Each item is allocated a unique, persistent identifier, a positive integer prefixed with the upper-case letter Q[why?], known as a "QID". This enables the basic information required to identify the topic that the item covers to be translated without favouring any language [35].

2.1.11 Smart Flanders

Smart Flanders 2.0 [36] is an Initiative of the Flemish government for the digital transition from local authorities to smart cities.

VLOCA: Vlaamse Open City Architectuur (Flemish Open City Architecture) [37] is an initiative to streamline all local Flemish smart city initiatives. Th architecture of VLOCA is depicted in Figure 5. The Flemish government is promoting Open Data. In tender open data is the default and when not followed, the reason needs to be explained. Data follows the DCAT standard.

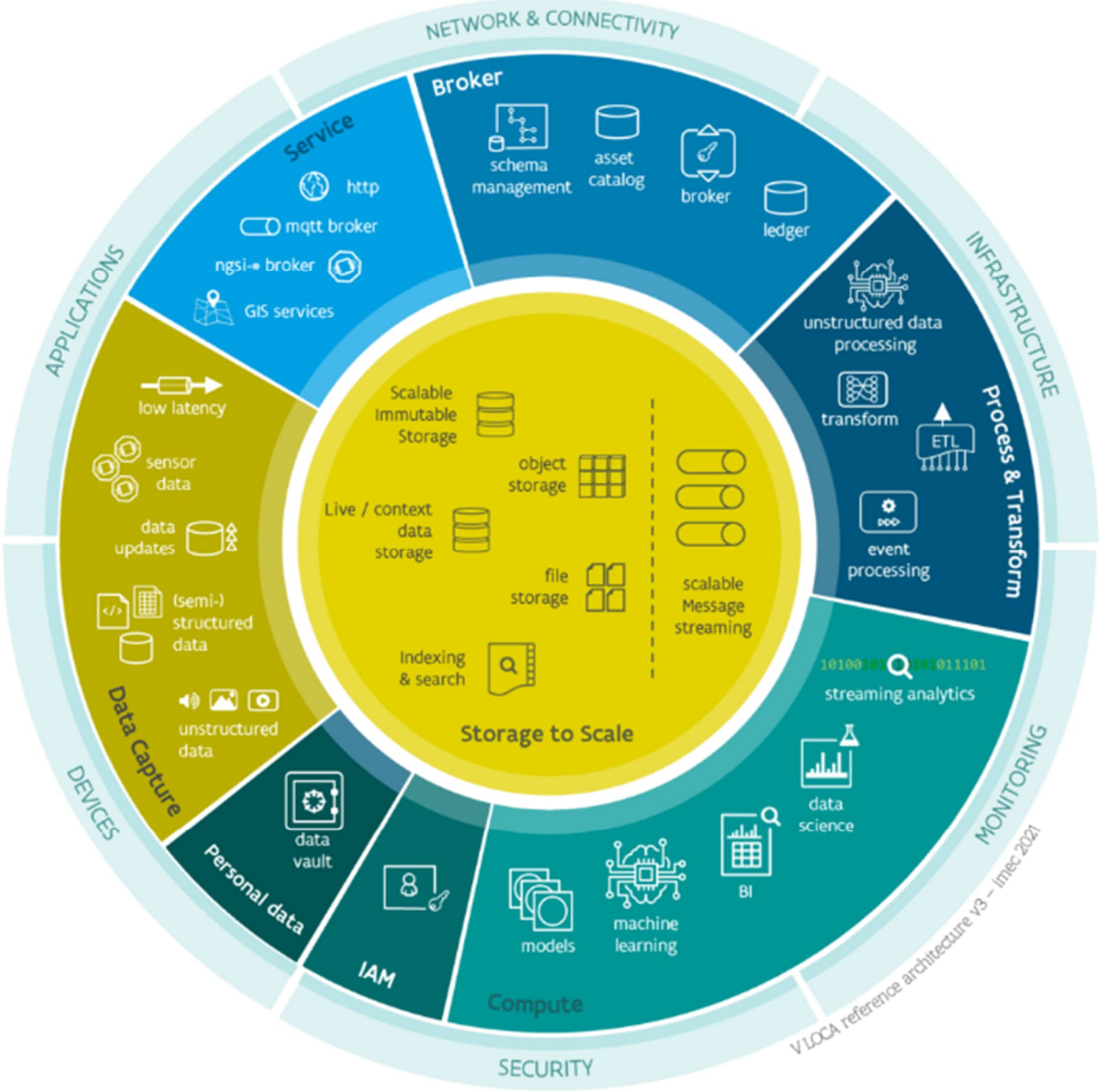

Figure 5. VLOCA Architecture.

2.1.12 Het Vlaams Datanutsbedrijf

“The Flemish Data Utility Company wants to stimulate citizens’ trust in sharing data, by focusing on responsible and secure data sharing. At the same time, we want to give oxygen to the Flemish economy by making data more findable and exchangeable, and by building bridges between citizens, companies and associations for better cooperation. We are a neutral third partner and catalyst for innovative initiatives and we a stimulate economic and social prosperity.” [38, 39]. The targets of the Flemish economy are depicted in Figure 6. Architecture of the Flemish economy is provided in Figure 7.

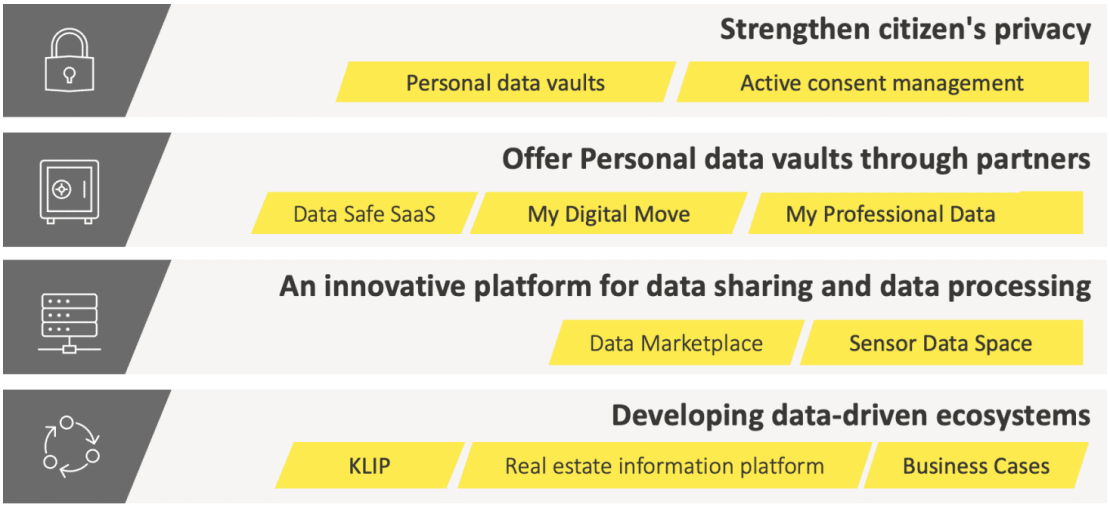

Figure 6 The targets of the Flemish economy.

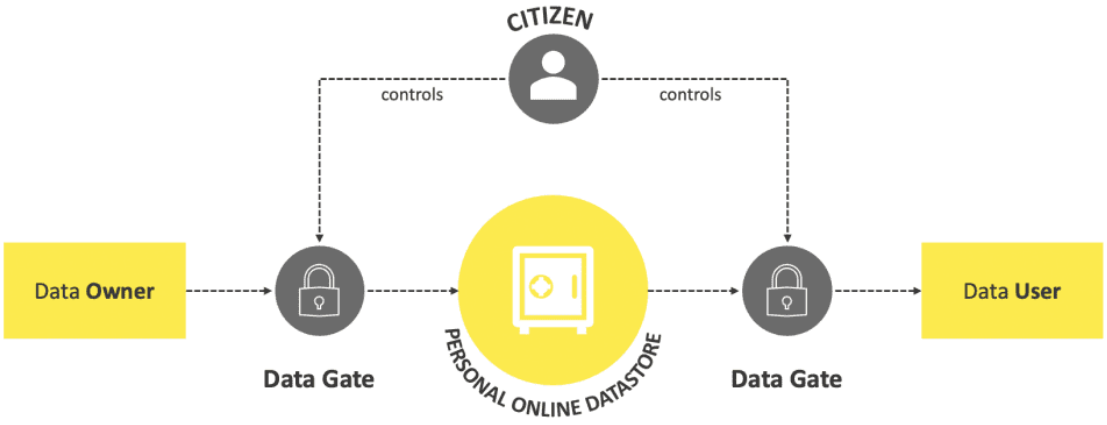

Figure 7 Architecture of the Flemish economy.

2.1.13 OPENDEI

OPENDEI [40] stands for "Open Data for European Interoperable Energy Services". It is a European Union-funded project, part of the European Commission's Horizon 2020, aimed at developing an open and interoperable data ecosystem for the energy sector. The project focuses on developing a framework that allows energy-related data to be shared and accessed by different stakeholders in the energy sector, such as energy producers, grid operators and consumers. The framework is designed to be flexible, scalable and secure, with the goal of enabling the development of innovative energy services that can benefit both consumers and energy providers. The project provides a set of tools and guidelines for managing and governing energy-related data including data quality assessment, data provenance tracking and access control policies.

2.2 Applications

When speaking about architectures and technologies it is always essential to consider what kind of application needs they try to fullfill. In addition, it is quite an essential challenge how to consider the logical and physical architectures in a relevant way. In order to give reader a view to the practical challenges in the selected sector specific application ares, we have opened in the section the main service areas which we have focused in this work: maintenance services, energy flexibility services, buildings related services, mobile services for consumers and traffic services.

2.2.1 Maintenance services

The maintenance of big buildings usually requires outsourcing of the building automation (monitoring and control) to be executed by a building automation company. Such an automation service company need to have access to the automation resources and information of buildings. Typically the service company need also collect referred information for making smart analysis and reasoning of problems in the buildings, otherwise they cannot make their automation services properly for their customers, who typically are owning organization of such buildings. However, the collected information is business sensitive, and therefore also critical for the owners and also valuable to the service company. Let's take an example, the need to change an HVAC filter is detected automatically via sensors, a request for tenders is created for other maintenance service companies to perform the change of the HVAC filter physically, Figure 8. The serviceperson who is going to make the operation need to have access rights for the spesific required operation. The state of the practise has been that the access rights are set separately for each service person manually by the admin of the automation service company. This has led to the problems in maintenance of the access rights, because this requires quite much manually created actions from the admin. This is because the access rights should be allowed only temporarily but not continuously, because the information and automation resources in a building are critical asset for the building maintenance service company and for the owners and tenants of the building.

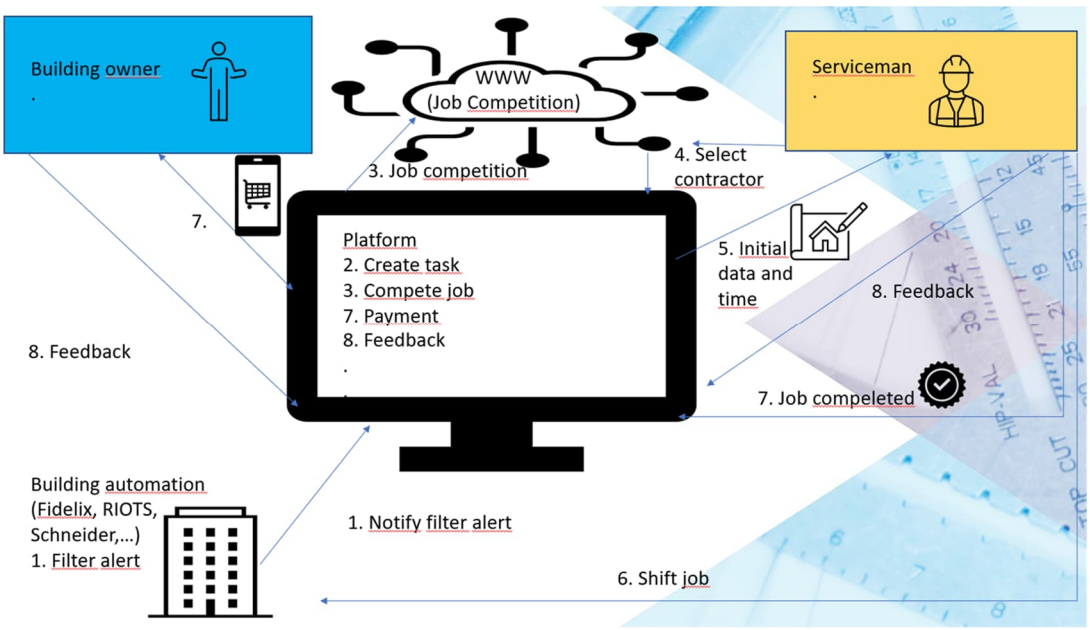

Figure 8. An example of maintenance services

2.2.2 Energy flexibility services

The current energy crisis in Europe highlights the need to use all the available energy resources in an efficient manner. Buidling new energy production resources is expensive and slow, and simultaneously the energy consumption is increasing quite rapidly due to the new needs coming e.g. from electric mobility needs. On the other hand, there are more and more new type of smaller scale energy production resources, such as solar power plants, which energy production capabilities are not properly under control of the energy stakeholders for natural reasons. The balance of consumption and production in the electric grid is very essential, and it today usually reached due to the automatic energy resources controlled by grid operator(s). However, achieving the balance can be very expensive, and it is visible during the peak hours of a day in the form of very high energy prices.

Therefore, it is estimated in the beginning of TIOCPS project that the capabilities to control the energy consumption and production during such peak hours could be beneficial. Such capabilities are here called as energy flexibility, referring to down or upscale energy consumption and production. Today, also EU is driving towards establishment for energy flexibility markets, and there are emerging energy flexibility marketplaces. Let's take an example related to management of energy sensitive resources of buildings, Figure 9. The buildings have automation devices to control the heating and cooling, which are currently working quite much under control of the device manufacturer and the local user. There are also automation service providers who provide more smart monitoring and control. However, there is quite much lack of energy flexibility aggregation functions that are able to interact with the energy market.

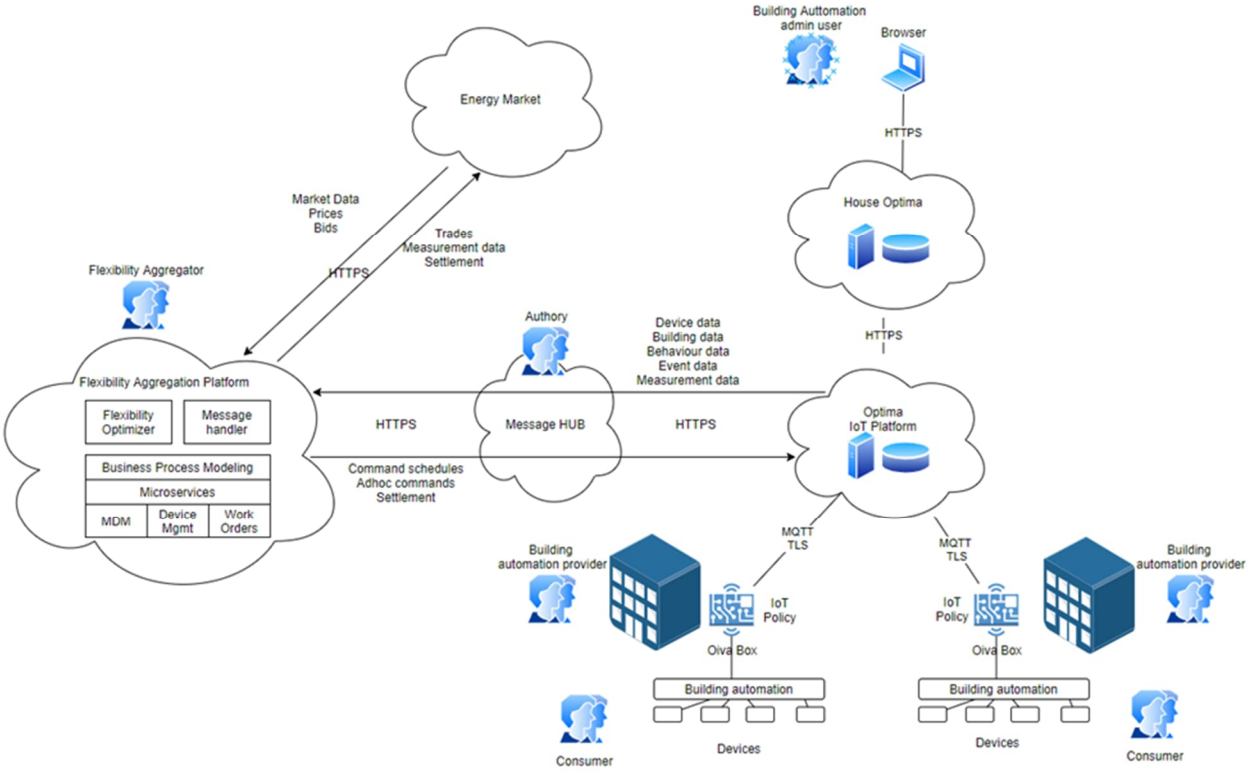

Figure 9. An example of energy flexibility services

In addition, when looking at the challenge from the consumers and prosumers point of view, it is important to get some monetary value or incentives, for making practical actions to lower consumption or increase

production. However, there are essential challenges related to scale and digital trust. The smallest tradeable unit in the energy flexibility market seems to be bigger than what individual consumer or prosumer is able to offer. In addition, the process for forecasting future production and consumption is quite a demanding process. Therefore, smart energy flexibility aggregation capabilities are needed, and the stakeholders owning energy flexibility resources need to combine their flexibilities in order to be involved in the energy flexibility markets, e.g. to establish energy communities. In addition, each of the referred stakeholders need to get monetary benefits or incentives in order to motivate them to offer their flexibilities into the energy flexibility markets.

2.2.3 Buildings related services

The buidings have usually a wide set of computing devices, which are manufactured by several independent players. Let's take an example related to a HVAC system of a factory, which has several sensors and controllable devices related to energy sensitive processes of a building, Figure 10. The HVAC system is at end nodes used for providing heating (boiler) and cooling (Chiller) services for the buildings internally. HVAC system consumes electrical energy for the utilisation of pumping motors and natural gas components (for heating and cooling respectively). Electricity measuring device is a telemetry device used for measuring the electricity consumption for offices and machinery. Control Units are the components used for controlling the HVAC infrastructure and the current available energy consumption data. Control units can be connected to Remote Terminal Units and Secure IoT Gateways by using MBUS protocols. The data collection layer presents services and tools for data collection. This layer is composed of a set of adapters, e.g. IoT adapters, external system adapters, API or open data adapters, The low-level hardware-based security is tackled by Secure IoT gateways. These gateways encrypt any data generated at end nodes, in cooperative way with the HSM at server side, and pass the secured data to some IoT platform. For example, Middleware and Context Broker is connected with the security layer that is used for communication with the real world (e.g. IoT sensor entities) and enables a contextual classification of data to provide monitoring and able to used for other services like elastic search.

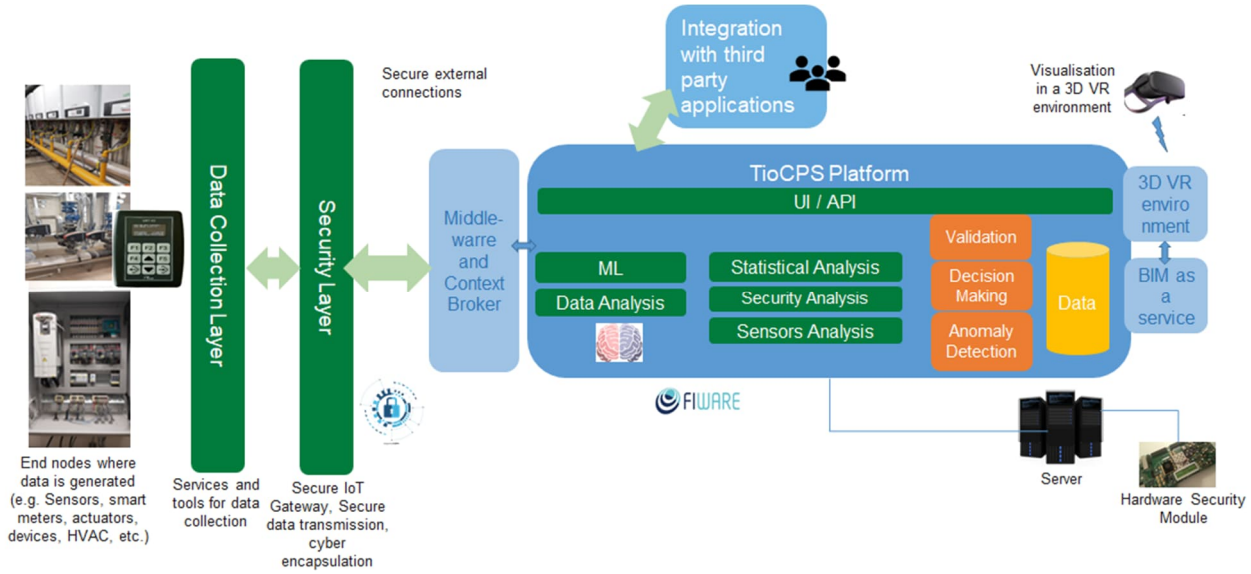

Figure 10. A example of buildings energy related services.

In the Figure 10 example, an IOT platform (TIoCPS platform in the Figure 10) presents a framework which enables the utilisation of services and tools (e.g. containers). This platform provides the BIM-as-a-Service which is used to model the targeted buildings and enable an information filtering and retrieval mechanism for energy management. A flexible ontology infrastructure is used to describe the taxonomy of the hardware components and any end node devices, connected peripherals, sensory data and their characteristics. The AI and ML tools are also integrated in the main platform. For instance, the decision-making module will be one of the main components that provides intelligent control of devices through solid and consistent decisions, powered by AI algorithms for energy saving. Anomaly Detection at System is planned to be deployed for analysing building conditions, building anomalies, faulty situations in the targeted smart building system. Validation of AI outputs helps the overall compliance with project goals and the results will be stored in a database. Since BIM enables 3D interpretation of the building data, a 3D Virtual Reality application is planned to be used for the visualisation of the infrastructure and its energy saving maps,

A state of the art view to the layered architecture can be defined in multiple ways, for example like depicted in Figure 11. Data acquisition layer is composed of adapters (IoT, external system, OpenData, API adapters) which acquire any input data from end nodes. Security layer enables end-to-end security which is capable of encrypting any transmitted data generated at end nodes. At server side an HSM is utilised to manage cryptographic operations. Middleware layer enables contextual harmonisation of data through a broker and present backend tools for advanced analytics, and data storage and persistence (including BIM-based data filtering, ontologies and semantic framework). Interoperability layer presents REST APIs and container management over a Service Oriented Architecture. Authentication layer is a vertical layer which enables the node and person authentication, authorisati and auditory actions. Interaction and Visualisation is at top layer which visualises the mahine learning and monitoring analysis results integrated with the system validation tools. The interfaces in the referred layered architecture can rely on the state of the art technologies, such as for example, IEEE 802.11F-2003 - Modbus/TCPIP, IEEE 1900.2/4/4a - Radio Frequency, ISO/IEC 20922 – MQTT, ISO 11898-3 CANBUS, IEC 62541 -OPC-UA, IEEE 802.15.4-2006 – BLE, IEEE 14575™-2000 – UART, ISO/IEC 15408 Common Criteria Standards (EAL4+), NIST-800-22 True Randomness Test criteria, and Industry Foundation Classes IFC2x Edition 3 (IFC 2x3).

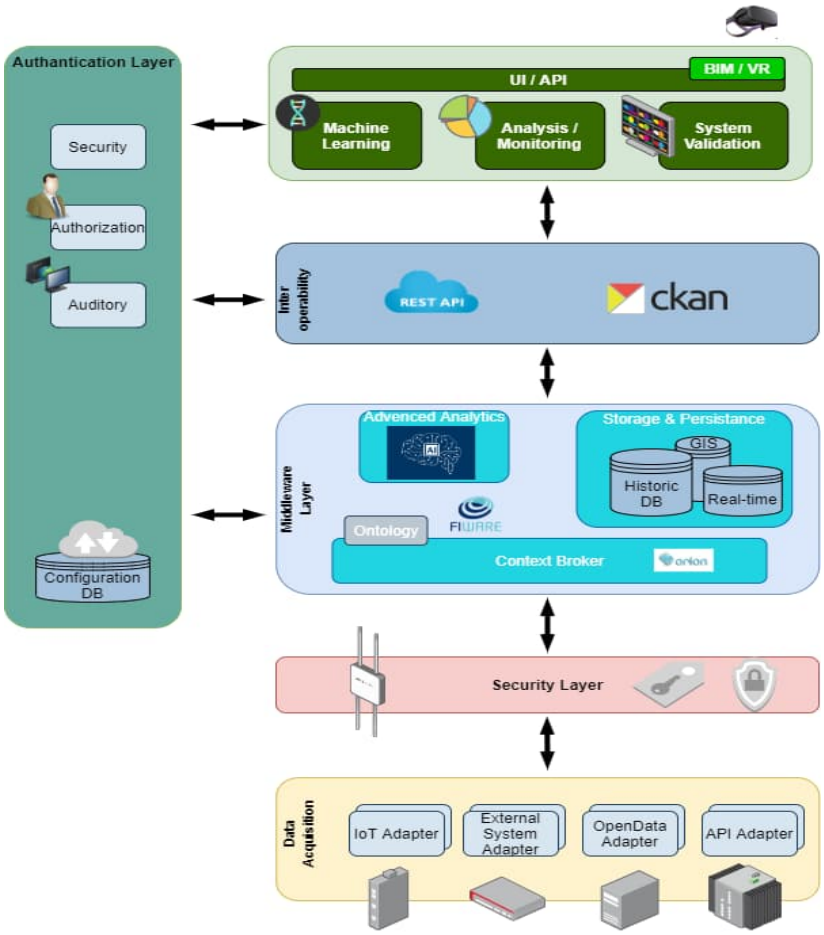

Figure 11. An example layered architecture for the management of smart buildings.

2.2.4 Mobile services for consumers

The mobile domain typically includes considering various kinds of wireless devices, on top of which specific services are provided for the owners and users of the referred devices by some specific service provider. The application field is typically quite a large, and therefore let's take an example called as hunting safety, Figure 12. Typically, the wireless devices such as e.g. smart watches, and connected with the specific service provider cloud, via which the owner of the device can apply advanced services relying on the information exposed from the device. This pattern seem to happen also with other types of devices, like e.g. hunting dog collars, which are used for tracking hunting dogs in the forest. It is obvious that when people are moving around in the same area in the forest with different aims, they are not aware about each other. Such a situation can be very dangerous, there is risk for hunting accidents, which are present when using the current technological practices.

When looking at such ad hoc wireless situations, there are several concerns. First of all, the information monitored by the referred devices is business sensitive for the referred service providers. Secondly, the information exposed from the devices is privacy sensitive for the owner/user of them. This especially true for the location information, because it may reveal presence and location of a person for unfriendly people.

There is a need to share privacy sensitive presence and location information between systems and users. There are technologies available for sharing such data, e.g. Bluetooth, which allow broadcasting and also connection type of channels for exchanging information in ad conditions in the forests. However, there is lack of solutions for digital trust and privacy in this kind of situations.

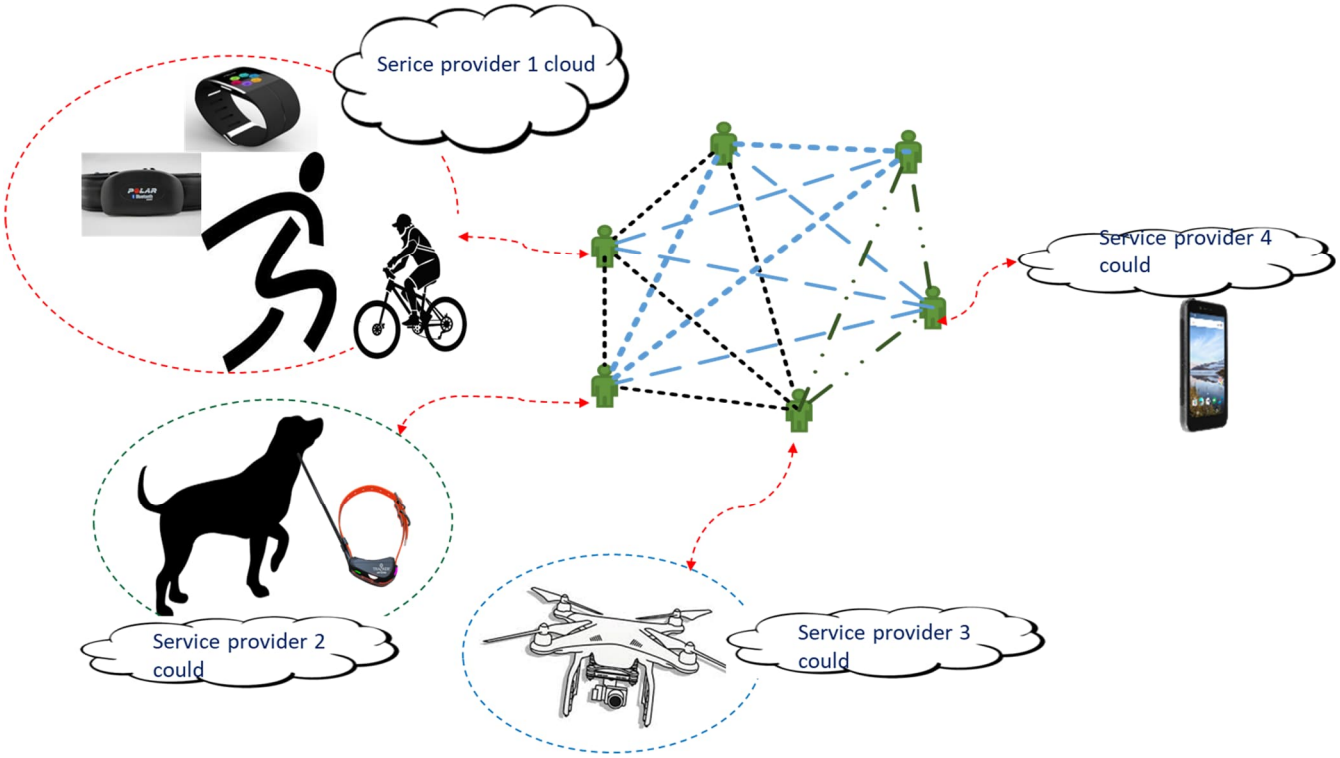

Figure 12. An example of mobile services for consumers.

2.2.5 Traffic services

An example of Traffic related service is depicted in the Figure 13 – services exposed from traffic cameras. These services typically follow speed of vehicles, take pictures from the vehicles that drives too fast, and organize speed penalty in such cases for the driver. Traditionally, this kind of action has been carried out based on the instantly recorded speed, but recently calculating average speed based on two cameras has been studied and taken into use. In addition, the roles of municipalities and cities have been increased in parallel with of the role of polices. The challenges have been related to the access and use of the referred recorded information of the traffic cameras. More specifically, authentication of the user when trying to access CCTV through SSH application, trying to use and share the recorded video information. This kind of situation includes also challenges related to privacy and trust, which is required between the authorities, municipalities/cities and the users.

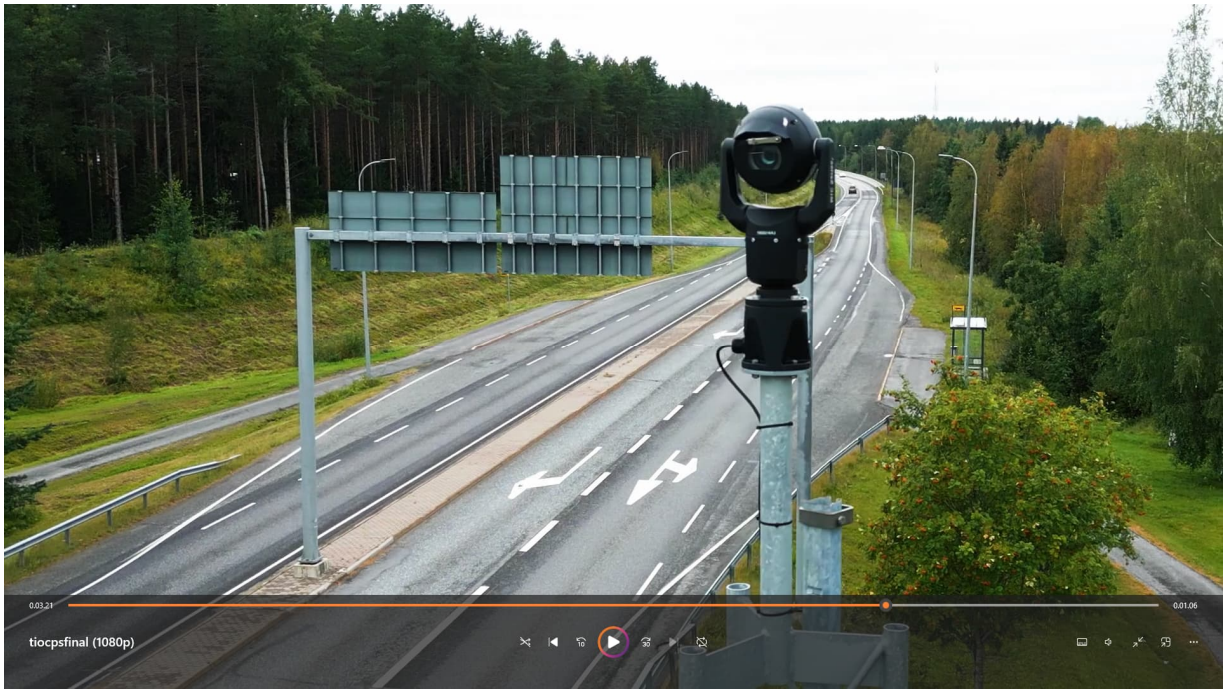

Figure 13. An example of a traffic related service.

2.3 A Discussion

It is estimated that the enterprises apply currently different kinds of IoT platforms for realizing the services exposed from the physical assets that they are hosting. The information sharing in current M2M/CPS systems is usually happening directly between the back-office systems of enterprises using APIs interfaces. There are several approaches for facilitating some kind of data sharing between such systems, like explained in chapter 2 of this deliverable. The GAIA-X, IDS, MyData and Energy Web approaches are shortly analysed in the following for looking at the potential architectural patterns which could be applicable for the TIoCPS concept.

GAIA-X approach try to avoid monolithic, centralized authorization and rather approach the problem via the use of decentralized, federated identifiers and services. It does not provide centralized data storage, either. Instead, each company decides for itself where its data is stored, as well as who may process it and for what purpose **Error! Reference source not found.**. The MyData approach leans towards decentralized operation, which suggests the use of self-sovereign identifiers (SSI) with technology support provided by operators to help users remain in control of their data. MyData approach to the authorization and permissions is to keep them separate from the data flow, and allocate verifying the operation to data operator. MyData model can have many such data operators. The reference architecture model of IDSA has IDS connectors as components for joining heterogeneous partner data sources to a common ecosystem. IDS relies on centralized certification authority where the trustworthiness of participants are ensured, which leads to a high level of trust but a heavy onboarding process relying on centralized identification, authentication, and authorization. Energy Web has not central authentication, and the solution is based on decentralized, blockchain-anchored identities, decentralized identifiers, verifiable claims, and smart contracts. User control over their data is the core principle of self-sovereign identities. With DIDs and verifiable claims, trust establishment is fine-grained and based on selective exposure of personal data. The application of self-sovereign seems to ease GDPR compliance and make data sharing

to third parties to be decided by the users. The concept of application registry with identities and verifiable claims in a way establish possibility to set-up a community for a specific geographical area of local energy grid. When comparing these existing solutions to the needs set-up by TIoCPS scenarios, there are several essential patterns that could be useful. The MyData model with decentralized operation looks to be quite well in-line with the need of owners and CPS service providers to be in control of their data. The Energy Web approach with decentralized, blockchain anchored identities, decentralized identifiers, verifiable claims, and smart contracts looks quite applicable for solving the requirements to support of GDPR.

The TIoCPS concept targets to enable trustworthy data sharing between heterogeneous enterprise systems with physical CPS assets, which may be owned by individuals or companies. The focus is mainly into the challenges related to authorization and access control, which solutions are especially needed in solving the innovation, business and technological challenges related to physical CPS resources -, vertical enterprises - and authorization zones, for making an essential breakthrough in trustworthy and smart information sharing with CPS applications. Based on the analysis of potential functional elements/components for the TIoCPS concept, some essential potential building blocks have been discussed in the following. Because of CPS systems are inherently multiple stakeholders' systems it is obvious that there is need to have some common reliable and trustworthy storage for the digital trust related agreements. A possible base for such solution seem to arise from recently developed technologies around blockchains, distributed ledger (DL) type of technologies. When any owner of the resources have decided something, and made an agreement with some other one, then such a resulting contract need to be stored so that it can be later found out in a reliable way. When any owner wants to withdraw from the contract, then it's decisions should be taken into use immediately by all the parties. This kind of situation is seen to be needed especially in the applications which apply privacy related issues. The owners and users of resources, back-office systems of enterprises, special 3rd party applications and various heterogeneous physical resources need to be able to interact with the referred shared digital trust base. Therefore, adaptable connectors from heterogeneous systems need to be provided by the TIoCPS concept. The connectors need to support e.g. registrations, negotiation of contracts, sign the contracts, checking status of contracts, token management to be used for ensuring trustworthy accesses. The aim is to help the communicating entities to exchange data directly between the endpoints, e.g. between the enterprises or between physical devices, so that the most recent status of contracts is taken into concern. This refers to one-to-one information sharing using the agreed endpoints. In addition, the system needs to support trust with one-to-many type of information sharing.

2.4 References

1. Alliance for Internet of Things Innovation. Online: https://ec.europa.eu/digital-single-market/en/alliance-internet-things-innovation-aioti accessed 21st Feb 2019.
2. IoT 2020: Smart and secure IoT platform. 181p. Available online (accessed 10th Jan 2018) http://www.iec.ch/whitepaper/pdf/iecWP-IoT2020-LR.pdf
3. ITU-T Study Group 13, Next Generation Networks – Frameworks and Functional Models: Overview of the Internet of Things, International Telecommunication Union, Geneva, 2012.
4. ZVEI – German Electrical and Electronic Manufacturers' Association, Industrie 4.0: The Reference Architectural Model Industrie 4.0 (RAMI 4.0), Frankfurt am Main, 2015.
5. Industrial Internet Consortium, Industrial Internet Reference Architecture (Version 1.7), Object Management Group, Needham, MA, US, 2015.
6. Internet of Things – Architecture Consortium, The IoT Architectural Reference Model (ARM) - D1.3, European Commission, Luxembourg, 2012

7. Fiware open specifications. Online documents (Accessed 13th Jan 2018). https://forge.fiware.org/plugins/mediawiki/wiki/fiware/index.php/Summary_of_FIWARE_Open_Specifications
8. SmartM2M Virtualized IoT architectures with Cloud Back-ends. ETSI TR 103.527 V0.2.1 (2018-01)
9. https://www.trusted-iot.org/
10. Amazon Web Services. Online documents (Accessed 13th Jan 2018). https://aws.amazon.com/documentation/iot/
11. Guth J., Breitenbucher U., Falkenthal M., Leumann F., Reinfurt L. Comparison of IoT Platform Architectures: A field study based on a reference architecture. CIoT'16.
12. IPSO Alliance, later OMA SpecWorks. Available online: https://www.omaspecworks.org/ (accessed on 21 Feb 2019)
13. The Internet Engineering Task Force (IETF). Available online: http://www.ietf.org/ (accessed on 21 Feb 2019)
14. ETSI M2M/Smart M2M. Available online: http://www.etsi.org/ (accessed on 21 Feb 2019).
15. One M2M forum. Available online: http://www.onem2m.org/ (accessed on 21 Feb 2019)
16. Latvakoski, J.; Alaya, M.B.; Ganem, H.; Jubeh, B.; Iivari, A.; Leguay, J.; Bosch, J.M.; Granqvist, N. Towards Horizontal Architecture for Autonomic M2M Service Networks. *Future Internet* 2014, *6*, 261-301.
17. [Online]. Available: https://gaia-x.eu/.
18. "Project GAIA-X: A Federated Data Infrastructure as the Cradle of a Vibrant European Ecosystem. German Federal Ministry for Economic Affairs and Energy (BMWi). October 2019 release.".
19. [Online]. Available: https://mydata.org/.
20. [Online]. Available: https://internationaldataspaces.org/we/gaia-x/.
21. [Online]. Available: https://international-data-spaces-association.github.io/DataspaceConnector/.
22. [Online]. Available: https://www.energyweb.org/.
23. [Online]. Available: https://energyweb.org/technology/applications/ew-switchboard/.
24. [Online]. Available: https://github.com/energywebfoundation/switchboard-dapp
25. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, European Commission
26. [Online]. Available: https://ec.europa.eu/commission/presscorner/detail/en/IP_20_2102
27. G. Cattaneo, G. Micheletti, M. Glennon, C. L. Croce and C. Mitta, The European Data Market Monitoring Tool, Publications Office of the European Union, 2020
28. J. Crémer, "Competition policy for the digital era," European Commission, 2019
29. [Online]. Available: https://www.fiware.org/
30. [Online]. Available: https://www.arrowhead.eu/
31. A. Botta, W. D. Donato, V. Persico and A. Pescapé, "Integration of cloud computing and Internet of Things: A survey," *Future Gener. Comput. Syst.,* vol. vol. 56, p. pp. 684–700, 2016
32. J. Delsing, IoT Automation: Arrowhead Framework, Boca Raton, FL, USA: CRC Press, 2017
33. [Online]. Available: www.flatturtle.com
34. [Online]. Available: www.wikidata.org
35. [Online]. Available: https://en.wikipedia.org/wiki/Wikidata
36. [Online]. Available: https://smart.flanders.be/.
37. [Online]. Available: https://vloca-kennishub.vlaanderen.be/.
38. [Online]. Available: https://www.vlaanderen.be/digitaal-vlaanderen/het-vlaams-datanutsbedrijf/the-flemish-data-utility-company
39. [Online]. Available: https://www.vlaanderen.be/digitaal-vlaanderen/het-vlaams-datanutsbedrijf
40. [Online]. Available: https://www.opendei.eu/
41. Juhani Latvakoski*, Marc Roelands, Marko Tilvis, Laura Genga, Gabriel Santos, Goreti Marreiros, Zita Vale, Lode Hoste, Wolfgang Van Raemdonck, Nicola Zannone. Horizontal Solutions for Cyber-Physical Systems evaluated in Energy Flexibility and Traffic Accident cases. Available online https://itea4.org/index.php/project/result/download/7344/Horizontal%20Solutions%20for%20Cyber-

Physical%20Systems%20evaluated%20in%20Energy%20Flexibility%20and%20Traffic%20Accident%20cases.pdf (accessed 28th Sep 2023)

3 State of the Art Analysis on CPS information level

We list the existing solutions for three different areas relevant to the building of a secure, trustworthy, and privacy-aware data sharing: (1) access control model, (2) secure data transfer protocols and (3) access control policy enforcement technique with data watermarking. We also highlight our contributions concerning how data or information is managed securely with trace and tracking ability in distributed environment.

3.1 Technologies

3.1.1 Access control models

There exist some standard access control models such as DAC, MAC, ORBAC and RBAC.

Discretionary Access Control (DAC) [1] is an access control model where restriction of access to objects is done based on the identity of subjects. The controls are discretionary in the sense that a subject with certain access permission is capable of passing that permission on to any other subjects. For example, an access control to files and folders in Unix system where user can define permission for other users to access their file or folder is a clear example of DAC. One implementation of DAC is the access control list that has been used widely in operating, networking, and database management system. Although the ability of passing access permission on to any other subjects seems to match with the requirements in smart traffic and another use cases in the scope of TiOCPS, DAC fails to fulfil other important requirements such as the ability to express complex permission assignment that involves purposes, obligations, and conditions, which are the most important elements for expressing the privacy-aware policy. Moreover, DAC cannot support data and role hierarchy expression. In addition, DAC has the problem of controlling the permission transfer and policy change, especially when data is shared to third party in the system. Another drawback of DAC is the ability to express the separation of duties.

Mandatory Access Control (MAC) [1] is used widely in the operating system, databases, and networking system. MAC refers to a type of access control model by which the system constrains the ability of a subject or requester to access or perform some sort of operation or action to an object or resource. In practice, the subject refers to an entity that can be a user or application; object refers to the files, directories, ports (in networking), or databases (tables or attributes in database management system). In MAC, subjects, and objects each have a set of security attributes and when a subject makes an attempt to access an object, an authorization rule enforced by the system examines the security attributes and then the decision can be made whether the access can take place. To determine if the operation on the object by a subject is allowed or not, those parameters will be tested against the set of the authorization rules made by the policy maker or administrator of the system. MAC provides the central control of the security. User or subject does not have the rights to assign or override the policy unlike ACL, which allows subject to make decision or override the access policy. MAC provides more control level compared to ACL as both subject and object carry the secured attributes that need to be checked or tested by system for every access attempt. The disadvantage of MAC lies in the complexity of the configuration, since for each resource (application, data) and subjects (user) must be determined, which access authorizations are necessary. This tends to be very difficult for the system that works with a large number of users and resources. Another down part is that MAC is not designed to enforce privacy policies and barely meet privacy protection requirements, particularly, purpose binding (i.e., data collected for one purpose should not be used for another purpose without user consent), conditions and obligations. The purposes and obligations are a part of the requirements for privacy-aware policy. Purposes are not only used to sharpen the access control but also to enforce the security protection.

Role-Based Access Control (RBAC) has been introduced in many research literatures [1][2][3]. RBAC restricts data access based on a person's role within an organization and has become one of the main methods for advanced access control. The roles in RBAC refer to the levels of access that employees have to the data. Using only user's role as condition for granting access to data is not sufficient for expressing a complex and fine-grain access control policy, particularly, in the contextual and privacy-preserving environment. For example, system that needs to differentiate access levels in the same role, and most importantly, system that

needs to express obligations or purposes. To complement this weakness, a context-aware and privacy-aware access control model was proposed [4]. In accordance with the spirit of the RBAC model, a privacy-aware RBAC (P-RBAC), access permission is granted to user based not only on user's role but also on the result of the occurrence events in the system or purpose of access. The context can be anything ranging from spatial, temporal to user pre-defined context. This new approach offers rich, fine-grain and flexible way to express the privacy-related policies. The authors realize the privacy-awareness by adding other entities to the core RBAC model such as conditions, obligations, and purposes of access. It is important to note that although standard RBAC cannot express the obligations, purposes, and conditions, it can provide many features that are necessary for expressing access control policy in smart traffic use case, such as the ability to express data and role hierarchy, permission transfer as well as separation of duties.

OrBAC (Organization Based Access Control) [5] is another access control model that the access permission is granted to user under a specific role in particular organization and contexts. OrBAC supports the control of data as well as user in system like organization structure and access permission is granted based on user's role in an organization. In addition, it can also express the access permission in contextual environment, role and data hierarchy, separation of duties as well as permission transfer.

ABAC (Attribute Based Access Control) [1] defines an access control model whereby access rights are granted to users through the use of policies which combine attributes together. ABAC is becoming well known and considered as a "next generation" authorisation model because it provides dynamic, fine-grained, context-aware, and intelligent access control. ABAC uses attributes as building blocks in a structured language that defines access control rules and describes access requests. Attributes are sets of labels or properties that can be used to describe all the entities that must be considered for authorisation purposes. The traditional ABAC model consists of 4 main entities (e.g., subject, action, resource and environment) where each entity may hold multiple attributes.

3.1.2 Data transfer protocols

Corporations, institutions, organizations, and other entities have daily requirements for file sharing, both within their organizations and among their larger external networks. These requirements are compounded by the need to demonstrate compliance with various industry mandates, such as the GDPR. File Transfer Protocol (FTP) is the go-to protocol for sending files. However, there also many highly secure protocols are more effective alternatives to transferring files that help to avoid the many risks associated with FTP. In this section, we list several industrial and widely used file transfer protocols.

a. SFTP

SFTP [6] allows organizations to move data over a Secure Shell (SSH) data stream, providing excellent security over its FTP cousin. SFTP's major selling point is its ability to prevent unauthorized access to sensitive information, including passwords, while data is in transit. The connection between the sender and receiver requires the user to be authenticated via a user ID and password, SSH keys, or a combination of the two. Because SFTP is more secure than FTP solution, it is often used for business trading partners to share information as it is platform independent and firewall friendly, only requiring one port number to initiate a session and transfer information.

b. FTPS

This protocol is known as FTP [6] over SSL/TLS, is another option for businesses to employ for internal and external file transfers.

FTPS has two security modes, implicit and explicit. Implicit requires the SSL connection to be created before any data transfer can begin. With Explicit SSL, the negotiation takes place between the sender and receiver to establish whether information will be encrypted or unencrypted. This means sensitive files or credentials can be set to require an encrypted connection before they will be shared. Like SFTP, the FTPS protocol can use a second factor authentication for added security.

c. AS2

Applicability Statement 2 (AS2) [6] is used to transfer Electronic Data Interchange (EDI) information in a secure way. AS2 wraps the data to be transferred in a secure TLS layer so it can travel from point to point over the internet with encryption as well as digital certifications for authentication. AS2/EDI is a transfer protocol in the retail industry, particularly with larger companies that require it for trading partner communications. This facilitates the efficient, secure, and reliable exchange of information and removes much of the chance for human error.

d. HTTPS

Hypertext Transfer Protocol Secure (HTTPS) [6] adds security to HTTP by offering certificate authentication. Additionally, it encrypts a website's inbound traffic and introduces an encryption layer via TLS to ensures data integrity and privacy. HTTPS protects a web visitor's identity and secures account details, payments, and other transactions involving sensitive details. When it comes to transferring files, this protocol enables the use of a simple but secure interface for uploading data from business partners or customers.

e. MFT (Managed File Transfer)

MFT [6] supports each of the secure FTP solution options listed above (SFTP, FTPS, AS2, and HTTPS) for secure data transmissions among internal users and external entities. This method includes an extensive list of security features that make it an ideal choice for meeting the stringent guidelines of many industry regulations.

MFT uses standards for GPG (GNU Privacy Guard) and PGP (Pretty Good Privacy (PGP) encryption to encrypt, sign, and decrypt files. It can also encrypt files automatically at rest in targeted folders. The ability to centralize your file transfers using MFT also gives you valuable reporting capabilities that display user access and all associated file transfers.

Not only does managed file transfer give you a rock-solid method of exchanging critical business information with vendors and trading partners securely, it also supports workflow automation, file transfer monitoring, notifications, and auditing. This means you can enhance productivity for your team in a variety of ways while keeping security at the forefront.

3.1.3 Techniques for secure sharing of data

In this section, we highlight the existing techniques for protecting data integrity.

3.1.3.1 Data watermarking

In this section, we present several watermarking techniques used for protecting ownership and preventing unauthorised tampering of electronic data.

1) Implementation of Embedding and Extracting Invisible Watermarking

The authors in [7] introduce a watermarking technique for protecting ownership and preventing unauthorized tampering of multimedia data (e.g., audio, video, image, and text). This technique can be used for Image authentication and to verify the originality of an image by detecting malicious manipulation; the ultimate goal to the watermark is retrieve the right owner information from the received data in a correct way. In this work, an image is taken from colour image (24 bits) type and from BMP file type and is converted into gray scale image (256 bits) and then converted into binary file by using one of filters (Sobel, Prewitt, Robert) to find edge detection of original file. Data storage process is performed in original image in edge points corresponding to the same place in a binary image. These edges are specified randomly based on location of the edge mod 3 and then specifying one of values (R, G, B) randomly to store data in it. As a result, invisible watermark is not noticeable

to viewer and without any degrade the quality of the content. The product invisible watermark is robust against distortions processes and resistant to intentional tampering solely intended to remove the watermark.

2) Blind Invisible Watermarking Technique in DT-CWT Domain Using Visual Cryptography

The authors in [8] present a method for digital image copyright protection by using a blind invisible and robust image watermarking scheme based on Dual Tree Complex Wavelet Transform (DT-CWT) and Visual Cryptography concept (VC). This method does not require that the watermark to be embedded into the original image which leaves the marked image equal to the original one. In the concealing and extracting process, the image is transformed in the complex wavelet domain to generate a secret and a public share respectively, using LL sub-band features and a VC codebook. To extract the watermark from the attacked image, the secret and public shares are stacked together. To improve the visual quality of the extracted watermark, a post process called reduction procedure is also proposed. The experimental results show that the proposed method can withstand several image processing attacks such as cropping, filtering and compression etc…

3) CryptMark: A Novel Secure Invisible Watermarking Technique for Color Images

The authors in [9] present a new technique for secure invisible watermarking technique for color images. Th novel method uses cryptography and watermarking methods simultaneously to provide a double layer protection to the digital media which can be an effective technique for Digital Rights Management (DRM) system. The proposed method securely hides binary information in colour image media, and securely extracts and authenticates it using a secret key. Experimental results prove that the proposed invisible watermarking techniques is resilient to 90% of the well-known benchmark attacks and hence a failsafe method for providing constant protection to the ownership rights.

4) An invisible watermarking technique for image verification

The authors in [10] propose a new method for invisibly watermarking high-quality color and gray-scale images. The proposed method is intended for use in image verification applications, where one is interested in knowing whether the content of an image has been altered since some earlier time, perhaps because of the act of a malicious party. It consists of both a watermark stamping process which embeds a watermark in a source image, and a watermark extraction process which extracts a watermark from a stamped image. The extracted watermark can be used to determine whether the image has been altered. The proposed technique is better than other invisible watermarking techniques for the verification application; these include a high degree of invisibility, colour preservation, ease of decoding, and a high degree of protection against retention of the watermark after unauthorized alterations.

3.1.3.2 Existing secure & privacy-aware data sharing & management frameworks

The authors in [12] propose a data-sharing system that, using only decentralized trust, (1) hides user identities from the server, and (2) allows users to detect server-side integrity violations. To achieve (1), Ghostor avoids keeping any per-user state at the server. To achieve (2), Ghostor develops a technique called verifiable anonymous history. By using a blockchain, publishing only a single hash to the blockchain for the entire system once every epoch.

3.1.3.3 Towards Secure and Decentralized Sharing of IoT Data

The authors in [13] propose a new framework named Sash. In Sash, the blockchain is used to store access control policies and take access control decisions. Therefore, both changes to policies and access requests are correctly enforced and publicly auditable. Further, in this framework, the identity-based encryption is used to cater for cryptography-enforced access control while minimizing the overhead to distribute decryption keys.

Sash is prototyped by using the FIWARE open source IoT platform and the Hyperledger Fabric framework as the blockchain back-end.

3.1.3.4 Trusted data sharing framework

Data sharing is a multi-disciplinary process which involves not only enabling technology, but also legal considerations. Concerns over trust and security hinder the mass sharing of data, despite the benefits that can be gained from leveraging large volumes and variety of data for analytics, including machine learning artificial intelligence. To this end, the authors in [14], develops a Trusted Data Sharing Framework, aiming to guide organisations through the data sharing journey and outline key considerations for organisations to consider when planning data sharing partnerships. It also provides an overview of the key areas in data sharing and helps users think through the entire process to structure their data sharing arrangements.

3.2 Applications

The proposed data access control and sharing model can be used in any application use cases such as, smart traffic, smart building, smart mobility, … Where data needs to be shared across systems and there is a need to control and trace & track who has accessed and used those data, the proposed solution helps.

This offline authentication scheme is the multi-factor authentication where none-time-based OTP (One Time Password) is used in combination with traditional username and password. In addition, Sirris & Macq are working on a new data sharing mechanism where data is attached with a pre-defined privacy-aware access control policy which will govern/control the way data is used and processed when it resides on data destination system out of control of data source system.

Offline multifactor authentication. There exist number of multifactor and one-time password authentication tools developed by well-known companies such as Google, Microsoft and AWS authenticator. However, those authenticators cannot be used in the context of Macq (smart traffic) use case and requirements. Macq requires to authenticate and trace user that access to the offline devices (e.g., cameras that are temporarily or permanently disconnected to the network) in a secure and trustworthy way. In addition, Macq wants to have centralise control of devices access and be able to track and trace who have accessed to devices without having access to devices' access log. Moreover, Macq wants to prevent even root/admin user from accessing the device without authorisation even if they know root/admin user password. The later requirement is important especially to prevent unauthorize access to device, for instance, in case Macq's employee left the company, but still have/know root password to access the devices and since some devices are not connected to the network, remote password's device update is not possible and frequent physical access to device for password update is possible, but not practical. The new and innovative authentication solution is required to address these special set of requirements.

To address those challenges, we propose an offline multifactor authentication mechanism that works without the need to have time synchronisation, which is not the case for the existing authentication method such as Google authenticator which is time-based. Our proposed solution is none time-based, and the secret values can be reused and still maintain security.

3.2.1 Secure data sharing in M3 system

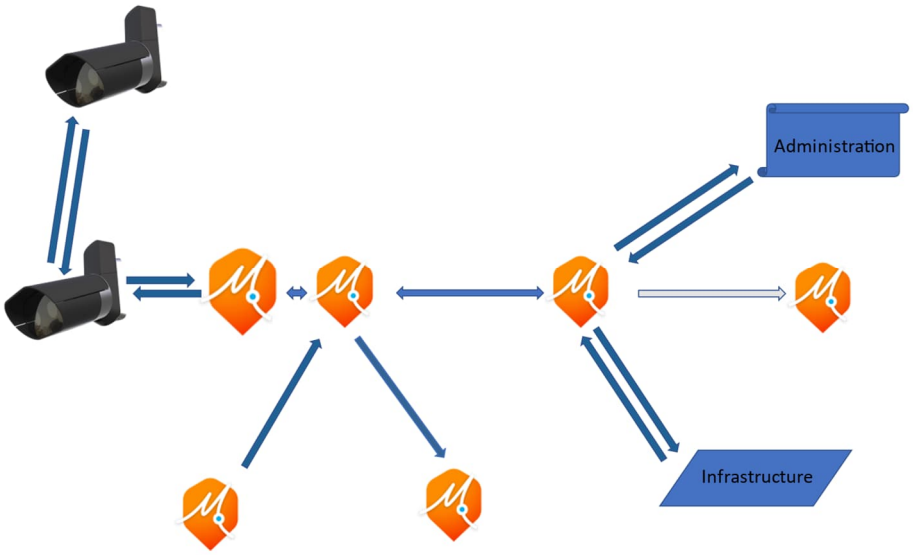

Figure 14. High-level architecture of data sharing in smart traffic use case.

The Macq QCAM – M3 ecosystem consists of several edge and back-end systems. The backend consists of several cascading and redundant M3 servers that connect to administrative servers and operational infrastructure servers or edge devices. The font-end is built with QCAM camera's and/or AI boxes.

In the scenario, there are following points of information exchange

- Between Edge and M3
- Between M3 and administrative servers
- Between M3 and operational infrastructure
- Between Edge devices

In D2.1 we introduced the super connection between camera and M3 with a number of features that this communication should be able to handle.

3.2.2 Health Checks & Monitoring

There is a fast and a slow process in this topic: status and health. Status indicates if hard- and software is functioning. A status can be ON, OFF or UNKNOWN. UNKNOWN covers the fact that there can be a cascading system of edge devices. From a status an alarm can be deduced. When hard- or software is failing this should be communicated directly.

Health indicates if the system and its components are functioning well. Degrading health must be detected to allow an early intervention. To detect a drop in performance and the performance needs to be monitored.

Data requirements:

- Both health and status are small amounts of data.
- Communication must avoid to many updates in case of a flickering status
- When the system goes off-line and becomes back online priority should be given to the current status

It must be possible to connect to existing monitoring and ticketing software.
This part of the protocol should not allow to intervene on the device.
To be decided how to configure the way status, health and alarms are calculated and communication settings.
In the functional analysis it was mentioned:

- Protocol should allow external commands (e.g. reboot, clear disk space, reinitialize secondary storage?)
    - Limited list of commands or free-form shell commands?
    - Should this be in the protocol or via SSH?

The safety requirements for this goal are clearly very different from check & monitoring. It is excluded from this topic.

3.2.3 Camera Software Updates from M³

Update User World software
Update Operating System (Linux)
Update hardware components (flash, eprom, ...)

Before starting an update, the system must be checked that all conditions are fulfilled to allow the update to reach the end. Ideally there must be a fallback mechanism when the update fails. If after a failing update there is the possibility of losing communication with the device the user launching the update must be warned at forehand.

3.2.4 Update Neural Networks from M³

Update from M3 to CAM seems covered by 'Camera Software Updateq from M3'.
Version of Neural Network is dedicated to the camera.
Trusthworthy by Watermarking of Neural Networks see NextPerception project
Privacy issues when downloading training data. What data is needed if training is divided between camera and server?

3.2.5 Back Up & Restore Camera Settings from M³

Used when a camera is replaced or reset to factory settings.
This implies identification of an installation, not the camera.
During the update there is physical access to the camera.
Access to the M3 server is through the network. The update protocol must be secured but allow the field technician to use it without physical access to the server.
During update it must be possible to exclude certain parameters that are invalidated by the camera replacement.

3.2.6 Easy First Camera Setup in M³

Avoid double configuration in camera and M3
Identification must be easaly communicatable over the phone or via SMS
Functionality for camera to 'announce' itself

- Should contain camera ID, GPS info?
- Admin user accepts/declines new camera on the server
- While not accepted, data does not get sent but stored locally so recovery can work

Smart Recovery of Missing Data

Observation: this means that the edge devices stores the data which can be a privacy issue. Some devices are sold in a context were storing data is explicity excluded.
There are cases when the Server wants same data again.
It must be possible to have multiple servers that request the same data.
Recovery must be guided by the server.
There must be a mechanisme to identify data.

3.2.7 Optimized Bandwidth Usage

- Prioritization on detections.
- Reserve enough bandwith for NTP.
- Images can be downloaded later.

3.2.8 Bulk Configuration from M³ on Installed Cameras

This is mainly an implementation issue in M3.
It does however impose some requirements for the protocols:

- Edge devices must be able to publish their parameters.
- Edge devices must be able to publish their software and protocol version(s)

3.2.9 Smart & Fast Data Delivery

- To reach the timing "blacklist hit gets 4 seconds from vehicle passing camera to being on M³ user web browser screen"
- Some customers will want data as soon as plate & country code OCR is finished
    - Even if no plate, customers are interested in the fact there was a vehicle
- Define when to send initial data, when to send any updates (everything in one go, or initial send after vehicle detected and 1 update after entire pipeline is finished, or multiple updates, …)
- How will recognition identification work? To indicate which recognition is being updated

3.2.10 Extend Camera from ANPR Data to Smart Data

From the functional analysis:

- Data types: recognitions, minute data, triggered data, statistics, … (counting, O/D on intersection, …)
- Protocol should allow sending/fetching of more than only individual detections
- Other types:
    - Aggregated data
        - Periodically
        - Only storing, data fetched by remote via API
    - Individual data
        - Event based: sent when occurring
        - Periodically (in bulk)
        - Only storing, data fetched by remote via API
        - Detections
            - Regular
            - Anonymized
                - Blurred images?
                - Anonymization information (salt ID to know which hashes can be compared, …)
            - Object paths (O/D on intersection, …)
    - Incident/Triggers (can involve multiple detections, e.g. collision)
        - Event based: sent when occurring
        - Periodically (in bulk)
        - Only storing, data fetched by remote via API

That is all true but the camera should not be a "one-trick pony". Smart Data also means other kind of recognitions. Most important is detection, tracking and prediction of Vulnerable Road Users behaviour. Arround the license plate there is also a vehicle with make, model, color, …
The 'Smart Data Protocol' must allow for object fusion.

3.2.11 Super Resolution on Data

Super resolution is a well know noition in vision technology (https://en.wikipedia.org/wiki/Super-resolution_imaging)
We should coin another term because here we something about the data that comes out of the image processing: Combine mulitple values over multiple recognitions to make the most probably recognition (or matching for speed or O/D, or black list, …)
The detections can also come from multiple edge devices.
The requirements for 'Super Resolution' seem already be covered by the 'Smart Data Protocol'

3.2.12 Schedule Video Recording/Download

To show the trustworthiness it is asked to have a video with a number of anotated detections that can be used to verify the good working of the cameras.

It must be possible to schedule multiple recordings of a video. Length of recording can be decided by duration or number of detections.
Detections must be releated to the video.
It must be possible to download the video.
The privacy policy of keeping this video on the camera must be guaranteed (How long is the video saved if it is saved on our server/camera)
A video is a huge amount of load. Where is it saved? Enough place? What about transfer of a video-file?

3.2.13 **Data Quality Checks**

The functional analysis is based on simple statistic taking into account variation due to known issues such as lightning conditions and traffic intensity
This can be considered a subtopic of the health check.
An AI implementation could be considered.
It should be possible to explain the bad or good quality of the data. Explainable AI if used.

3.2.14 Data sharing and access control system architecture

The architecture (see Figure below) consists of the Edge devices, Backends network (data source) and Backends network (data destination/third party) (for more detailed infos about the high-level system architecture see D2.1).

Access control policy definition and data preparation module is responsible for managing data access control policy and preparing data before sharing them to third party. This module can be deployed at Edge devices (if device has enough memory and processing power) or Backends network (data source).

Access control policy definition module is responsible for creating and managing data access control policy. Access control policy defines who can access data for what purpose, in which circumstances and in which conditions. Policy administration point is an interface allowing user to manage data access control policy. Access control policy is generated in “access control policy generator module”.

Once access control policy is created, the policy is binded with data. The data and policy are then packaged, which is encrypted with security key before sharing. At destination, only entity (or subject) having decryption key and being mentioned in access control policy (attached with data) can open and access data.

Below is the description of data flow (see Figure below):

1) through policy administration point (PAP) user creates access control policy using access control policy generator (ACPG) module
2) user defines data access conditions and
3) purpose and
4) ACPG generates access control policy based on those inputs.
5) access control policy is ready to is sent to data and access control policy binding module, to be binded with data
6) data and policy are packaged and encrypted with trusted encryption key.
7) the encrypted (data+policy) is ready to be shared
8) the prepared shared data is accessible through policy administration point.

Data access control and enforcement module is responsible for controlling access to data in accordance with the defined access control policy. As shown in figure below, this module takes “encrypted (data + policy)”

as an input. When user requests access to data, the encrypted (data + policy) is decrypted. The module extracts data and access control policy. However, data is not yet made available to user. User is granted access to data if and only if access control policy is evaluated by “policy decision point” with positive response. Positive response indicates that user is the right one, the purpose of access mentioned by user matches to purpose defined in access control policy and all conditions mentioned in access control policy are satisfied.

Below is the description of data flow (see Figure below):

1) user requests access to data through “policy enforcement point (PEP)"
2) PEP forwards request to “policy decision point (PDP)" to check if permission can be granted
3) PDP sends request to “data and access control extraction (DACE)" module to retrieve access control policy and shared data
4) encrypted data +policy is provided to DACE
5) Decryption key is provided to DACE
6) with decryption key, policy is extracted and
7) shared data is also extracted
8) access control policy is provided as input to PDP
9) PDP evaluates the access control policy and if it is positive, access is granted. PEP informs user
10) user is allowed to access shared data

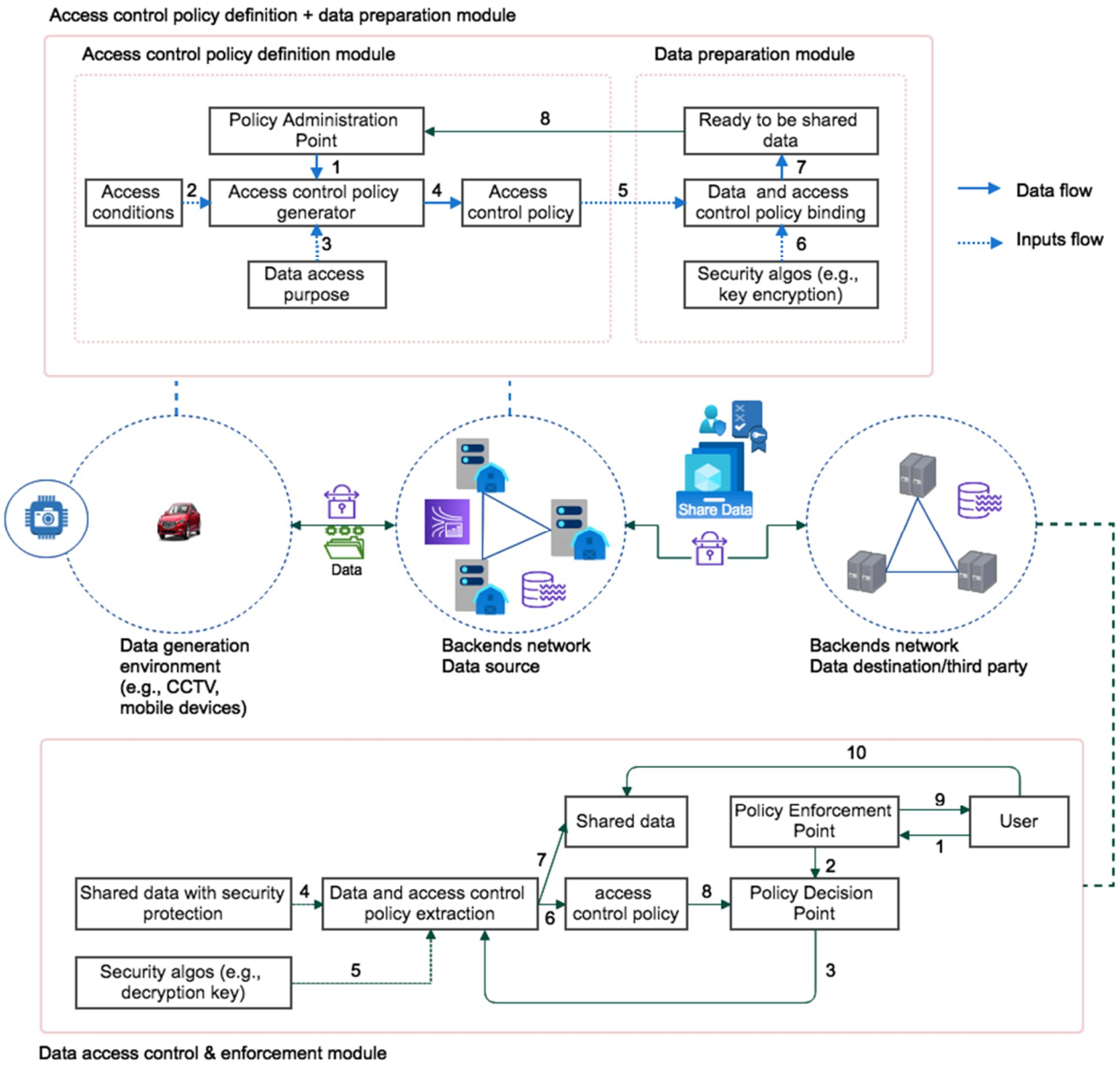

Figure 15. Data sharing access control system architecture.

Standard security protocols:

Each camera has the Macq CA certificate pre-installed as trusted.
This certificate is used for the HTTPS connections on at least the camera endpoints but can be used for all HTTPS connections.

Standard access control policy expression language & engine

- JSON
- AES for policy and shared data encryption
- TLS for secure communication

- MQTT for messaging protocol

3.3 Discussion

Contributions in TioCPS: Improvement of the existing data sharing model. The existing data access control models do not address the issue of miss using data after the access permission is granted. This means that abuse of the data usage is possible. Data can be shared to third party without data owner's consent and knowledge once access permission is granted. Our proposed access control and data sharing management takes into account the data usage, especially when data needs to be shared in the distributed environment. In the proposed data sharing scheme (see also the deliverable 3.3), data is attached with access control policy and access/usage log and protected by encryption key (in our proposed solution, the encrypted package of data and access control policy is shared instead of only data). The key used to encrypt the data and access control policy package is derived from two keys: (1) a key which is known to user, and (2) another key which is known to application only. In this way, the data is linked to application and to user who knows one of the keys. If the data is shared with unauthorised third party, it cannot be open even though the key that is known to user is exposed. Another important feature of the proposed data access control and sharing model is its ability to provide track and trace of data access and usage in the data processing workflow (e.g., if data needs to be shared between different entities in the system).

Concerning data integrity, in the proposed data sharing model, we use both visible and invisible watermarking technique to embed the identity of the organisation and company which is the destination where data needs to be processed. This allows for easy track and trace in case of data breach occurs. Moreover, the identity of the company is also encoded in the access control policy, binding the data to the policy and identity of company or organising that processes it.

3.4 References

[1]. Vincent C. Hu, David F. Ferraiolo, and D. Rick Kuhn. Assessment of Access Control System. National Institute of Standards and Technology, September 2006. https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7316.pdf

[2]. D.F.Ferraiolo, R.Sandhu, S.Gavrila, D.R.Kuhn, and R.Chandramouli. Proposed NIST Standard for Role-Based Access Control. ACM Transactions on Information and System Security, August 2001, pp.4(3):222-274.

[3]. Lorenzo D. Martin, Qun Ni, Dan Lin, and Elisa Bertin. Multi- domain and Privacy-aware Role Based Access Control in e- Health. Second IEEE International Conference on Pervasive Computing Technologies for Healthcare, Jan-Feb 2008, pp. 131-134.

[4]. Qun Ni, Bertino, Elisa, Lobo, Jorge, Brodie, Carolyn, Karat, Clare-Marie, Karat, John, Trombeta, and Alberto. Privacy- aware Role-Based Access Control. ACM Transaction Infor- mation and System Security, July, 2010, pp.24-3.

[5]. A.Bou, R. Baida, P.Balbiani, S.Benferhat, F.cuppens, and Y.Deswarte. Organization Based Access Control Model. 4th IEEE International Workshop on Policies for Distributed Sys- tems and Networks, June, 2003.

[6]. File transfer protocols. https://www.goanywhere.com/blog/2019/02/21/five-secure-file-transfer-alternatives-to-ftp

[7]. Kareem Mohammed Jubor. Implementation of Embedding and Extracting Watermarking. Online ISSN: 0975-4172 & Print ISSN: 0975-4350. Online ISSN: 0975-4172 & Print ISSN: 0975-4350.

[8]. Benyoussef M., Mabtoul S., El Marraki M., Aboutajdine D. (2013) Blind Invisible Watermarking Technique in DT-CWT Domain Using Visual Cryptography. In: Petrosino A. (eds) Image Analysis and Processing – ICIAP 2013. ICIAP 2013. Lecture Notes in Computer Science, vol 8156. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-41181-6_82

[9]. S. P. Mohanty, R. Sheth, A. Pinto and M. Chandy, "CryptMark: A Novel Secure Invisible Watermarking Technique for Color Images," 2007 IEEE International Symposium on Consumer Electronics, 2007, pp. 1-6, doi: 10.1109/ISCE.2007.4382120.

[10]. M. M. Yeung and F. Mintzer, "An invisible watermarking technique for image verification," Proceedings of International Conference on Image Processing, 1997, pp. 680-683 vol.2, doi: 10.1109/ICIP.1997.638587.

[11]. Qinlong Huang, Licheng Wang, Yixian Yang, "Secure and Privacy-Preserving Data Sharing and Collaboration in Mobile Healthcare Social Networks of Smart Cities", Security and Communication Networks, vol. 2017, Article ID 6426495, 12 pages, 2017. https://doi.org/10.1155/2017/6426495.

[12]. Yuncong Hu, Sam Kumar, Raluca Ada Popa. Ghostor: Toward a Secure Data-Sharing System from Decentralized Trust. In Ranjita Bhagwan, George Porter, editors, 17th USENIX Symposium on Networked Systems Design and Implementation, NSDI 2020, Santa Clara, CA, USA, February 25-27, 2020. pages 851-877, USENIX Association, 2020.

[13]. H. T. T. Truong, M. Almeida, G. Karame and C. Soriente, "Towards Secure and Decentralized Sharing of IoT Data," 2019 IEEE International Conference on Blockchain (Blockchain), 2019, pp. 176-183, doi: 10.1109/Blockchain.2019.00031.

[14]. Trusted Data Sharing Framework. https://www.imda.gov.sg/-/media/Imda/Files/Programme/AI-Data-Innovation/Trusted-Data-Sharing-Framework.pdf

4 State-of-the-Art Analysis on CPS communications

This chapter discusses the existing technologies, standards, frameworks and systems that are expected to be relevant to the different M2M service aspects to be covered by the layer.

4.1 Technologies

4.1.1 Trust and interoperability technologies researched within the TioCPS project

As one of the TioCPS project's aim is to research trustworthy, smart and interoperable information/data sharing, a few more novel technologies have been researched for the project demonstrators. The Google's Macaroon Tokens have been researched within the Serviceman scenario, for providing ad-hoc, delegated and restricted access in a secure way. Also, the Open API Specification (OAS) has been researched within the same use case, in order to find a way for different systems to be able communicate with each other.

4.1.1.1 Google's Macaroon Token

Macaroons are authorization credentials that provide flexible support for controlled sharing in decentralized, distributed systems. Macaroons are widely applicable since they are a form of bearer credentials, much like commonly used cookies on the Web, and have an efficient construction based on keyed cryptographic message digests.

Macaroons allow authority to be delegated between protection domains with both attenuation and contextual confinement. For this, each Macaroon contains caveats, i.e., restrictions, which are predicates that restrict the macaroon's authority, as well as the context in which it may be successfully used.

For example, such restrictions may attenuate a Macaroon by limiting what objects and what actions it permits, or contextually confine it by requiring additional evidence, such as third-party signatures, or by restricting when, from where, or in what other observable context it may be used. [19]

Usability

As there is no full specification/standard to how the caveats, i.e., restrictions are to be defined/notated within a Macaroon token, each implementor can arbitrarily decide how their Macaroon tokens' restrictions are annotated. This of course is not different from any other current web tokens, such as JWT i.e., JSON Web Tokens, technologies but renders the desired "universal" usability of the Macaroon as access tokens lacking in the context of trustworthy and interoperable systems.

One possible way to enhance that lacking aspect could be to use similar more standardized way of declaring the restrictions, such as what the public key infrastructure certificates have regarding e.g., validity time etc. As a comparison, the JWT tokens do have around seven standardized "claims" (i.e., restrictions) fields, but in addition to those, rest of the token fields are custom and service specific.

Security

HMAC-based macaroons suffer in one important way compared to public-key-based credentials: only the target service that originally mints a macaroon can check its validity. Specifically, the verification of an HMAC-based macaroon requires knowledge of its root key, and since this key confers the ability to arbitrarily modify the macaroon and its caveats, it cannot be widely shared. [19]

In the TioCPS project Serviceman scenario, this is acceptable, since all the tokens are created and verified by the Bittium Proxy Server, and the Proxy Server then communicates with the target service. The main downside of this is the fact that the original receiver of the token cannot create sub tokens in an offline manner, but only via the Proxy Server provided dynamic web UI, i.e., online connectivity is required to be able to create sub tokens, i.e., to be able to mint more restricted Macaroon tokens.

Also, within the TioCPS project Serviceman scenario, the Macaroon tokens are distributed in serialized form within QR code, that can be distributed also in an out-of-band manner, instead of sending the token over Internet to the end-user device. This can in some use cases be beneficial for added security.

4.1.1.2 Open API Specification

The OpenAPI Specification (OAS) defines a standard, programming language-agnostic interface description for HTTP APIs, which allows both humans and computers to discover and understand the capabilities of a service without requiring access to source code, additional documentation, or inspection of network traffic.

When properly defined via OpenAPI, a consumer can understand and interact with the remote service with a minimal amount of implementation logic. Similar to what interface descriptions have done for lower-level programming, the OpenAPI Specification removes guesswork in calling a service. An OpenAPI document that conforms to the OpenAPI Specification is itself a JSON object, which may be represented either in JSON or YAML format. [20]

Nowadays, web development frameworks such as FastAPI can create OAS documentation automatically during the development and build process of the web application, which makes it easy and fast to deliver the API specification of the service under development to the possibly already existing consumer applications, or consumer applications being under development simultaneously.

In the TioCPS project Serviceman scenario, the OAS was researched as a mechanism to make the Bittium Proxy Server to as automatically as possible to understand the target service API, so that when new target services are added, there would be little to no need at all to add additional modules/code into the Proxy Server.

4.1.2 Communication Technologies and Standards

4.1.2.1 Bluetooth

Bluetooth (BT) (https://www.bluetooth.com/) is a prevalent wireless standard that facilitates secure and user-friendly connections between devices within a short range [1]. It is especially apt for audio communication, stereo streaming, and managing sensors and devices in home environments. The introduction of Bluetooth Low Energy (BLE) has broadened the scope of this technology, enabling the integration of inexpensive, energy-efficient devices into the network. However, it's crucial to acknowledge certain limitations of this technology. For example, BT systems might necessitate internet gateways for data transmission to distant locations. Moreover, the protocol's simplicity could potentially lead to security threats such as eavesdropping and Denial of Service attacks. Despite considerable enhancements to the BT/BLE protocols over time, it's worth noting that potential security vulnerabilities might persist due to optional or selectable security features [2]. In summary, while Bluetooth technology provides substantial benefits in terms of convenience and cost-effectiveness, it's vital to thoroughly evaluate the potential risks and constraints associated with its usage. By implementing appropriate measures to address these issues, both businesses and consumers can continue to reap the numerous benefits of this pivotal wireless standard..

4.1.2.2 ISA100.11A

In industrial settings, the integration of wireless sensing and communication can present several obstacles, especially when updates are needed in less than 1 millisecond. The ISA100.11a (https://isa100wci.org/) standard was established to tackle this problem, providing a dependable and secure wireless alternative for non-critical monitoring and control tasks. This standard, which is built on IEEE 802.15.4, includes a variety of network components such as the SecurityManager, System Manager, GWs, backbone routers, and field devices [3]. Although the ISA100.11a standard can incorporate devices that use different communication protocols, it can also add a significant level of complexity that may hinder full device interoperability.

4.1.2.3 Near-Field Communication

Near-Field Communication (NFC) (http://nearfieldcommunication.org/) is a protocol that allows two electronic devices to connect at low speeds within a range of 10 centimeters or less [4]. Its simplicity and ease of use make it a popular choice for initiating more complex wireless connections. However, the limited range of NFC does not guarantee secure communications and can be susceptible to unauthorized access and data alterations [5]. To mitigate these security risks, some applications employ advanced cryptographic protocols to create a secure channel. Interestingly, current System-on-Chips (SoCs) incorporate security measures at higher layers, facilitating secure NFC implementations suitable for sensitive applications like contactless payment. In summary, while NFC provides user-friendly and straightforward connectivity, it's crucial to be aware of potential security threats and implement necessary precautions to ensure secure data transmission.

4.1.2.4 Wi-Fi

The IEEE 802.11 (https://www.ieee802.org/11/) standards, which form the backbone of Wireless Local Area Networks (WLANs), offer enhanced transmission, range, and throughput capabilities compared to the IEEE 802.15.4 standards. Wi-Fi, a widely used technology built on these standards, enables wireless connectivity to the Internet or corporate networks for consumer electronics.

Despite initial concerns about high energy consumption making it unsuitable for Cyber-Physical Systems (CPS) with wireless sensors, the development of low-power Wi-Fi devices has changed the landscape. These devices leverage existing infrastructure, well-known protocols, native IP-network compatibility, and a wealth of network management tools and knowledge. The IEEE 802.11 communication protocol is commonly used in IoT devices [6].

Wi-Fi-enabled sensors can implement standard security measures such as WEP, WPA/TKIP-PSK, and WPA2/AES-PSK to ensure data confidentiality, authentication, and availability as provided by the 802.11 standard.

Most Wi-Fi-based sensing solutions use infrastructure networks where sensors communicate directly with the Gateway (GW) or Wi-Fi Access Point (AP). This eliminates the need for additional GWs when connecting to the Internet. Furthermore, when sensors use internet protocols like UDP, there's no need for applications that encapsulate data for transmission to the Internet.

4.1.2.5 WirelessHART

The WirelessHART standard is an attractive choice for those in need of a wireless network that is not only easy to set up but also offers flexible installation alternatives and reliable communication [7]. This standard, which is built on IEEE802.15.4, includes a variety of elements such as gateways, network and security managers, field devices, adapters, and handheld devices. It is especially beneficial for mission-critical applications due to its ability to provide dependable end-to-end communication and manage end-to-end delay.

Both the ISA100.11a and WirelessHART networks are characterized by their low power usage, which allows wireless devices to function for prolonged periods on long-lasting batteries. In addition, they both prioritize security, utilizing AES-128 based encryption to ensure confidentiality and multiple layers of protection.

4.1.2.6 ZigBee

The IEEE 802.15.4 (https://www.ieee802.org/15/pub/TG4.html) standard provides an economical and energy-efficient solution for relatively low data rates [8]. It accommodates both peer-to-peer and star topologies and distinguishes between two types of devices: Full-Function Devices (FFDs) and Reduced-Function Devices (RFDs). ZigBee, an extension of the IEEE 802.15.4 standard, enhances this framework by incorporating mesh networking, making it a suitable choice for wireless connectivity applications such as home automation, and monitoring and control systems. However, it's important to consider that ZigBee might not be the optimal choice for applications with strict latency and reliability demands. The energy-efficient design of ZigBee and other IEEE 802.15.4 based solutions contributes to their popularity in sensor network applications [9]. Nevertheless, these solutions necessitate the use of gateways for internet data transmission. Additionally, ZigBee employs the Advanced Encryption Standard (AES) algorithm in counter mode, which can lead to significant code and time overheads, potentially posing challenges for nodes with limited resources.

4.1.2.7 GSM / GPRS

The General Packet Radio Service (GPRS) is a standard for mobile data that functions on 2G and 3G cellular networks [10]. It was conceived by the European Telecommunications Standards Institute (ETSI) as an advancement over previous packet-switched cellular technologies such as CDPD and i-mode. Presently, it is maintained by the 3rd Generation Partnership Project.

Unlike circuit-switched data, which bills per minute of connection time, GPRS is typically charged based on the total volume of data transferred during a billing cycle. If a user exceeds their GPRS plan's data cap, they may face additional charges per MB of data, experience reduced speed, or in some cases, be barred from using the service.

Despite the emergence of newer technologies like LTE CAT M1 and Nb-IoT that are set to supersede GSM/GPRS, studying automatic network switches in extreme network conditions remains relevant. This is because GPRS continues to serve as a backup channel in certain countries.

4.1.2.8 Nb-IoT

The Narrowband Internet of Things (NB-IoT) is a radio technology standard that facilitates a wide array of cellular devices and services [11]. This technology is designed with a focus on indoor coverage, affordability, extended battery life, and high-density connectivity [12]. The NB-IoT specification was officially established in LTE Advanced Pro in June 2016.

NB-IoT operates using a subset of the LTE standard, limiting the bandwidth to a single narrow-band of 200kHz. It employs OFDM modulation for downlink communication and SC-FDMA for uplink communications.

NB-IoT is an optimal solution for IoT applications necessitating frequent communications. It operates without duty cycle limitations on the licensed spectrum, making it highly efficient and reliable.

4.1.2.9 LTE CAT M1

LTE-M, also known as LTE-MTC (Machine Type Communication), is a low power wide area network (LPWAN) radio technology standard. It is designed to support a wide variety of cellular devices and services, with a particular emphasis on machine-to-machine and Internet of Things applications.

The LTE-M technology encompasses eMTC (enhanced Machine Type Communication) and was introduced in the 3GPP Release 13 (LTE Advanced Pro) in June 2016. It offers several advantages over NB-IoT, including higher data rates, the ability to transmit voice over the network, and mobility.

However, these benefits come at the cost of increased bandwidth and energy consumption, which makes LTE-M less sensitive compared to NB-IoT. Despite these challenges, LTE-M remains a promising technology for driving the growth of the Internet of Things and facilitating machine-to-machine communication.

4.1.2.10 LoRa and LoRaWAN

LoRa (https://lora-alliance.org/), an acronym for Long Range, is a cutting-edge technology designed for extended-range transmissions while maintaining low power usage [13], [14]. It's built upon spread spectrum modulation methodologies that are a product of chirp spread spectrum (CSS) technology. The inception of this technology took place in Grenoble, France by Cycleo, which was later acquired by Semtech.

LoRa operates on license-exempt sub-gigahertz radio frequency bands such as 433 MHz, 868 MHz (Europe), 915 MHz (Australia and North America), 865 MHz to 867 MHz (India), and 923 MHz (Asia). The data transmission rates with LoRa can vary from 0.3 kbit/s to 27 kbit/s, contingent on the spreading factor.

LoRaWAN, standing for Long Range Wide Area Network, represents the higher layers of this technology and functions in synergy with LoRa.

4.1.2.11 LTE Direct (Device to Device)

Introduced in the Release 12 specification, LTE Direct is a groundbreaking protocol that enables direct communication between nearby LTE devices [15]. This revolutionary communication method offers several significant benefits. It optimizes spectrum usage, boosts throughput, and enhances energy efficiency [16]. Furthermore, Device-to-Device (D2D) communication paves the way for location-based peer-to-peer applications and services, marking it as a transformative technology in the mobile technology landscape.

4.1.2.12 Free bands and protocols, 151-169 MHz, 430 MHz and 868/915MHz free bands

A variety of frequency bands, including the 151-169 MHz, 430 MHz, and 868/915MHz free bands, are utilized across different systems. These bands are compatible with both proprietary and open standards. Cutting-edge technology can be custom-designed to meet specific conditions by leveraging these free frequency bands. This technology can be crafted by integrating the best practices from existing standards, proprietary standards, and incorporating research findings.

4.1.3 Communication Protocols and Encryption

4.1.3.1 Internet Protocol

The Internet Protocol (IP) is a fundamental component of the Internet protocol suite, facilitating the transmission of datagrams across diverse network boundaries. Its routing functionality is vital for creating and sustaining internetworking. The IP is engineered to deliver packets based on the IP addresses in the packet headers, from the originating host to the target host. It outlines packet structures that encapsulate the data to be delivered and addressing methods that tag the datagram with source and destination details.

However, during the early Internet's design phase, security issues were not sufficiently foreseen, leading to numerous vulnerabilities in many Internet protocols. These vulnerabilities have been underscored by network attacks and subsequent security evaluations.

The IP offers only best-effort delivery and is deemed unreliable due to its connectionless nature, unlike connection-oriented communication. This can lead to various fault conditions such as data corruption, packet loss, and duplication. As routing is dynamic, each packet is treated individually, and different packets may be routed to the same destination via different paths, resulting in out-of-order delivery to the recipient.

To mitigate these issues, all network fault conditions must be identified and rectified by the participating end nodes. The upper layer protocols of the IP suite are tasked with resolving reliability issues, such as buffering network data to ensure correct sequencing before delivery to an application.

IPv4 incorporates safeguards to ensure that an IP packet's header is error-free. In contrast, IPv6 operates without header checksums because current link layer technology provides adequate error detection.

4.1.3.2 Transmission Control Protocol

The Transmission Control Protocol, or TCP, is a vital part of the Internet Protocol suite. It ensures that data is transmitted in a reliable and orderly manner between different applications on separate hosts via an IP network. Often paired with IP, it's commonly known as TCP/IP.

TCP forms the backbone of key internet applications such as email, file transfer, remote administration, and the World Wide Web (WWW). It's an integral part of the Transport Layer within the TCP/IP suite. Secure Sockets Layer/Transport Layer Security (SSL/TLS) frequently operates over TCP.

TCP is a connection-oriented protocol, meaning a connection must be established between the client and server before data can be transmitted. The server needs to be ready and listening for connection requests from clients to establish this connection.

While TCP's three-way handshake and error-checking mechanisms boost reliability, they can also lead to increased latency. For applications that prioritize speed over reliability, the User Datagram Protocol (UDP) may be a better fit. UDP provides a connectionless datagram service.

Although TCP uses network congestion avoidance, it's not completely immune to vulnerabilities. There are several known TCP attacks, including Denial of Service (DoS), connection hijacking, TCP veto, and reset attack.

4.1.3.3 User Datagram Protocol

The User Datagram Protocol (UDP) is a crucial part of the Internet protocol suite. It enables computer applications to send datagrams to other hosts on an Internet Protocol (IP) network without needing prior communication. UDP uses a simple connectionless communication model with minimal protocol mechanisms. It provides checksums for data integrity and port numbers for addressing different functions at the datagram's source and destination.

However, UDP does not have handshaking dialogues, which can expose the user's program to the underlying network's unreliability. It does not guarantee delivery, ordering, or protection against duplicates.

UDP is suitable for applications where error checking and correction are either not needed or are handled in the application itself.

Time-sensitive applications often use UDP because dropping packets is preferable to waiting for delayed packets due to retransmission, which may not be possible in a real-time system. However, because UDP lacks reliability mechanisms, applications using it must be ready to accept some packet loss, reordering, errors, or duplication.

Most UDP applications do not use reliability mechanisms and might even be impeded by them. Applications such as streaming media, real-time multiplayer games, and voice over IP (VoIP) often use UDP. In these cases, packet loss is usually not a critical issue.

Many key Internet applications use UDP, including the Domain Name System (DNS), Simple Network Management Protocol (SNMP), Routing Information Protocol (RIP), and Dynamic Host Configuration Protocol (DHCP). Voice and video traffic are primarily transmitted using UDP, and real-time video and audio streaming protocols are designed to handle occasional lost packets.

Some VPN systems like OpenVPN may use UDP and perform error checking at the application level while providing reliable connections. The development of quality of service solutions is considered crucial by some businesses as both real-time and business applications are essential.

4.1.3.4 Message Queuing Telemetry Transport

The Message Queuing Telemetry Transport (MQTT) (https://mqtt.org/) protocol is a compelling choice for those in search of a network protocol that is lightweight, secure, and user-friendly, facilitating efficient message exchange between devices [17]. It functions solely over TCP/IP but can be utilized in conjunction with any other protocol that ensures ordered, lossless, and bi-directional connections. This characteristic renders it an optimal selection for remote connections where network bandwidth is at a premium or a minimal code footprint is required. Importantly, MQTT offers secure operations through mechanisms such as authentication, access control lists, role-based access control, and TLS and X509 certificate and OAuth authentication. These security features ensure the protection of your data even when it is being transmitted over a potentially insecure network.

4.1.3.5 Internet Protocol Security

Internet Protocol Security (IPsec) is a robust network protocol suite frequently employed in Virtual Private Networks (VPNs). It facilitates secure, encrypted communication between two computers over an Internet Protocol network. The suite includes protocols for establishing mutual authentication between agents at the beginning of a session and negotiating cryptographic keys for use during the session.

IPsec can protect data flows between two hosts (host-to-host), two security gateways (network-to-network), or a security gateway and a host (network-to-host). It offers a range of cryptographic security services, such as network-level peer authentication, data-origin authentication, data integrity, data confidentiality (encryption), and replay protection.

What sets IPsec apart is its operation at the IP layer, enabling it to automatically secure applications at that layer. This contrasts with other security systems like Transport Layer Security (TLS) and Secure Shell (SSH), which operate at the Transport Layer and Application layer, respectively.

While the original IPv4 suite had limited security provisions, IPsec provides an end-to-end security scheme at layer 3 of the OSI model or internet layer. This helps protect data and ensure secure communication.

4.1.3.6 QUIC

The latest iteration of HTTP, HTTP3, stands out for its use of QUIC, a departure from previous HTTPS versions that relied on a TCP and TLS combination for reliability and security. HTTP3 streamlines the connection process by using a single handshake, unlike the separate handshakes required by TCP and TLS, thereby reducing the time needed to establish a connection.

QUIC (https://www.rfc-editor.org/info/rfc9000) has gained widespread acceptance, with over half of all connections from Chrome to Google's servers using it [18]. It's also supported by other major browsers like Microsoft Edge, Firefox, and Apple Safari. QUIC aims to improve the performance of web applications that currently use TCP by establishing multiple multiplexed connections between two endpoints via UDP. It's poised to replace TCP at the network layer for many applications.

In addition to reducing connection and transport latency, QUIC also estimates bandwidth in each direction to avoid congestion. The protocol can be further enhanced with forward error correction, which is anticipated to boost performance in case of errors. This is seen as the next step in QUIC's evolution.

4.1.3.7 Advanced Encryption Standard

The Advanced Encryption Standard (AES), established by the United States National Institute of Standards and Technology (NIST) in 2001, is a protocol for encrypting electronic data. It's derived from the Rijndael block, a cipher family with different key and block sizes. For AES, NIST chose three Rijndael family members, all with a 128-bit block size but varying key lengths of 128, 192, and 256 bits. The U.S. government's adoption of AES speaks to its robustness and security. As a symmetric-key algorithm, AES uses the same key for both data encryption and decryption. It's part of the International Organization for Standardization / International Electrotechnical Commission (ISO/IEC) 18033-3 standard and is commonly found in various encryption packages. Moreover, AES is the only publicly accessible cipher that the U.S. National Security Agency (NSA) approves for top-secret information when used in an NSA-approved cryptographic module, highlighting AES's reliability as a trusted encryption standar.

4.1.3.8 Elliptic Curve Digital Signature Algorithm

The Elliptic Curve Digital Signature Algorithm (ECDSA) is a variant of the Digital Signature Algorithm (DSA) that leverages the principles of elliptic curve cryptography. It's important to highlight that, similar to other elliptic-curve cryptography methods, the public key's bit size required for ECDSA is roughly twice the security level, measured in bits. For example, if we consider a security level of 80 bits (implying an attacker would need to perform approximately 2^80 operations to discover the private key), an ECDSA private key would be 160 bits in size. In contrast, a DSA private key would need to be at least 1024 bits.

4.1.3.9 Rivest-Shamir-Adleman

The RSA cryptosystem, a renowned public-key encryption method established in 1977, employs a pair of large prime numbers to create a public key and an auxiliary value for secure data transmission. This public key is capable of encrypting messages, but decryption necessitates knowledge of the prime numbers. The security of RSA hinges on the factoring problem, making it challenging for attackers to crack the encryption. Despite its efficacy, RSA is relatively slow and is commonly used to transmit shared keys for symmetric key cryptography. The security strength of RSA is measured in bits, with a 112-bit public key offering 112 bits of security. In contrast, the ECDSA algorithm can achieve the same security level with just a 224-bit public key. This characteristic makes smaller key sizes more desirable for devices with limited resources, as it reduces the bandwidth needed for secure connections.

4.1.3.10 Transport Layer Security

The Transport Layer Security (TLS) is a vital protocol that safeguards communication across computer networks. It is widely used in various applications, including email, instant messaging, and Voice over IP (VoIP), and is most recognized for its role as the security layer in HTTPS. The main goal of TLS is to ensure privacy and data integrity between two or more communicating computer applications.

TLS operates through two layers: the TLS record and the TLS handshake protocols. These layers allow client-server applications to communicate over a network without the risk of eavesdropping or tampering. To establish a connection, the client and server participate in a handshaking procedure that negotiates a stateful connection. During this handshake, they agree on several parameters that determine the security of the connection, such as the encryption algorithm and cryptographic keys.

Once the connection is established, data transmission is encrypted using a symmetric-key algorithm. The keys for this encryption are uniquely generated for each connection and are based on a shared secret negotiated at the start of the session. This shared secret is secure and reliable as it cannot be intercepted or obtained by an attacker who positions themselves in the middle of the connection.

TLS also facilitates the authentication of the communicating parties' identities using public-key cryptography and certificates. Certificates are crucial to TLS security as they set and define the security level. This authentication is mandatory for the server and optional for the client.

The connection is also reliable because each transmitted message includes a message integrity check using a message authentication code to prevent undetected loss or alteration of data during transmission.

TLS can be configured to offer additional privacy-related properties such as forward secrecy, ensuring that any future disclosure of encryption keys cannot decrypt any past TLS communications. By carefully configuring TLS, it can provide enhanced security and privacy for sensitive information communication.

4.2 Applications

The trust and interoperability technologies researched within the TioCPS project were utilized within the TioCPS Serviceman scenario, by using the Google's Macaroon technology as access tokens to grant more granularized access to Building Automation System. The technology is envisioned to be used with some of the Bittium products/applications as well, for example to be able to grant time-limited and more granular access to e.g., analytics data for different customers and stakeholders.

Based on the research done within the TioCPS project, the Macaroon technology as such can be used for multiple access token related applications, and they can be taken into use within implementations in quite straight-forward manner. However, due to the restriction format not being standardized, the content of tokens remains arbitrarily decided by the implementer, which presents a further research problem for the CPS community, if the tokens are envisioned to be used as more general-purpose access tokens.

4.3 Discussion

TIoCPS project have made significant strides in the field of communications, with a particular focus on cyber-physical systems heterogeneity and data trustworthiness. These advancements of TIoCPS have paved the way for seamless connectivity between heterogeneous networks and enhanced security and trustworthiness for cyber-physical systems. One of the key technological findings was that the Macaroon tokens enable a good starting point for creating access tokens that can be distributed and further delegated. In combination with a Proxy Server (or a proxy library, if desired to be more integrated with a target service), the researched technology provided a novel way to grant more granular access to target service and to be able to further delegate the access in a secure and trustworthy manner.

The evaluated results of TIoCPS clearly prove that researched and implemented solutions are capable of fulfilling defined requirements of trustworthy and cyber secure communication solutions and therefore can be said that set goals regarding have been achieved, proving the feasible existence of secure and trustworthy cyber-physical environments. The developments of TIoCPS, regarding communications, allowed the deployment and test of the solution in four different application case studies. Nonetheless, each of the evaluated solutions includes findings which need to be taken into account when continued to improve solutions maturity up to the product level (e.g., technical optimization for the solutions and applications and from interoperability point of view standardization will be required).

4.4 References

[1] S. Zeadally, F. Siddiqui, and Z. Baig, "25 years of bluetooth technology," Future Internet, vol. 11, no. 9, 2019, doi: 10.3390/fi11090194.

[2] M. Cäsar, T. Pawelke, J. Steffan, and G. Terhorst, "A survey on Bluetooth Low Energy security and privacy," Computer Networks, vol. 205. 2022. doi: 10.1016/j.comnet.2021.108712.

[3] M. Kashef and R. Candell, "Performance of an ISA100.11a Industrial Wireless Network with Wi-Fi Interference", doi: 10.6028/NIST.IR.8239.

[4] Z. Cao et al., "Near-field communication sensors," Sensors (Switzerland), vol. 19, no. 18. 2019. doi: 10.3390/s19183947.

[5] A. Alrawais, "Security Issues in Near Field Communications (NFC)," International Journal of Advanced Computer Science and Applications, vol. 11, no. 11, pp. 621–628, 2020, doi: 10.14569/IJACSA.2020.0111176.

[6] É. Morin, M. Maman, R. Guizzetti, and A. Duda, "Comparison of the Device Lifetime in Wireless Networks for the Internet of Things," IEEE Access, vol. 5, 2017, doi: 10.1109/ACCESS.2017.2688279.

[7] P. A. M. Devan, F. A. Hussin, R. Ibrahim, K. Bingi, and F. A. Khanday, “A Survey on the Application of WirelessHART for Industrial Process Monitoring and Control,” Sensors 2021, Vol. 21, Page 4951, vol. 21, no. 15, p. 4951, Jul. 2021, doi: 10.3390/S21154951.
[8] Vaishali, Varsha A M, Tejaswini G, and Vandan M Shetty, “ZigBee Technology,” International Journal of Advanced Research in Science, Communication and Technology, 2022, doi: 10.48175/ijarsct-7036.
[9] K. EL GHOLAMI, Y. MALEH, and I. F. E. FATANI, “The ieee 802.15.4 standard in industrial applications: A survey,” Journal of Theoretical and Applied Information Technology, vol. 99, no. 15. 2021.
[10] A. Amin and M. N. A. Khan, “A Survey of GSM Technology to Control Remote Devices,” International Journal of u- and e-Service, Science and Technology, vol. 7, no. 6, 2014, doi: 10.14257/ijunesst.2014.7.6.14.
[11] E. M. Migabo, K. D. Djouani, and A. M. Kurien, “The Narrowband Internet of Things (NB-IoT) Resources Management Performance State of Art, Challenges, and Opportunities,” IEEE Access, vol. 8, 2020, doi: 10.1109/ACCESS.2020.2995938.
[12] V. Kumar, R. K. Jha, and S. Jain, “NB-IoT Security: A Survey,” Wirel Pers Commun, vol. 113, no. 4, 2020, doi: 10.1007/s11277-020-07346-7.
[13] Z. Sun, H. Yang, K. Liu, Z. Yin, Z. Li, and W. Xu, “Recent Advances in LoRa: A Comprehensive Survey,” ACM Trans Sens Netw, vol. 18, no. 4, 2022, doi: 10.1145/3543856.
[14] M. A. M. Almuhaya, W. A. Jabbar, N. Sulaiman, and S. Abdulmalek, “A Survey on LoRaWAN Technology: Recent Trends, Opportunities, Simulation Tools and Future Directions,” Electronics (Switzerland), vol. 11, no. 1. 2022. doi: 10.3390/electronics11010164.
[15] S. Kumar Gupta, J. Yusuf Khan, and D. Trong Ngo, “An LTE-Direct-Based Communication System for Safety Services in Vehicular Networks,” in Moving Broadband Mobile Communications Forward - Intelligent Technologies for 5G and Beyond, 2021. doi: 10.5772/intechopen.91948.
[16] S. Mumtaz, H. Lundqvist, K. M. S. Huq, J. Rodriguez, and A. Radwan, “Smart Direct-LTE communication: An energy saving perspective,” Ad Hoc Networks, vol. 13, no. PART B, 2014, doi: 10.1016/j.adhoc.2013.08.008.
[17] B. Mishra and A. Kertesz, “The use of MQTT in M2M and IoT systems: A survey,” IEEE Access, vol. 8, 2020, doi: 10.1109/ACCESS.2020.3035849.
[18] J. Iyengar and M. Thomson, “RFC 9000: QUIC: A UDP-Based Multiplexed and Secure Transport,” 2021, Accessed: Sep. 25, 2023. [Online]. Available: https://www.rfc-editor.org/info/rfc9000
[19] A. Birgisson, J. Gibbs Politz, U. Erlingsson, A. Taly, M. Vrable, M. Lentczner, “Macaroons: Cookies with Contextual Caveats for Decentralized Authorization in the Cloud”, Network and Distributed System Security Symposium, Internet Society, 2014
[20] Open Api Specification v.3.1.0, https://spec.openapis.org/oas/v3.1.0, 2021

5 State-of-the-art analysis on CPS devices, trust and security

Threat modelling enables discovery of potential threats and security mitigation that a given software system may be susceptible to. In essence, modelling tools and frameworks detail the ideas behind the system, which includes a catalogue of potential threats, alongside the parties that could exploit the threats, and their methods to do so.Threats discovered during the lifecycle of a system can be very expensive to fix; so, threat modelling tools and frameworks are used to identify threats and minimize the cost to fix by discovering them early in the development cycle. Based on the needs of a project or software different modelling methods can be used. Each are somewhat specialized for different parts of the process and more than one method can be used to fully model all threats.

5.1 Technologies

5.1.1 HW Security for Data Sharing - HSM

Hardware Security Module (HSM-PRIGM) and Secure Gateway (SGW) to enable data sharing over secure communication. As depicted in Figure 1, firstly, TCP/IP socket connections between Front-end node and HSM, and HSM and SGW should be established respectively. The Front-end node consists of end-node applications developed in Unity or web-based applications for visualizing the BIM models and updating synchronized IoT information obtained from CPS assets. The HSM is integrated to the Front-end node to manage the incoming and outgoing messages. Also, the HSM is integrated in the Cloud to manage the incoming and outgoing messages' security. On the other hand, in order to receive data from CPS assets, a bunch of sensing devices should be registered on Secure Gateway side where it manages the incoming and outgoing messages.

Both HSM and SGW are handling communication on their own sides, while the communication between them is encrypted. When these operations are concluded, the user should be able to request data from the CPS assets.

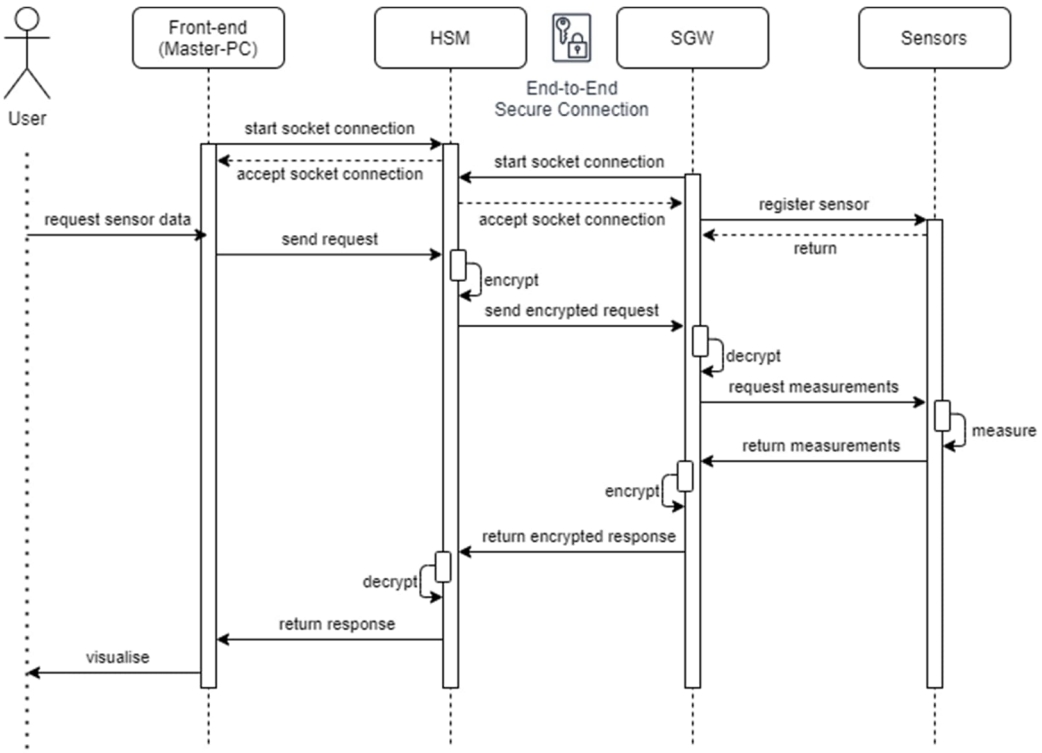

Figure 11. End-to-End Holistic Cyber-Physical Security

5.1.2 HW Security for Data Sharing - SoftHSM

In Figure 2, HSMs are used on Client (Unity) and Cloud side for encrypted data exchange over internet. This approach enables to use HSM functionalities into applications. Note that, since the HSM is costly device, a software version of the HSM is developed, namely SoftHSM, in order to use on the client applications. Also, the system described in Section 6.1.1 is used for receiving IoT data from the SGW in order to visualize the real-time data on Digital Twin application.

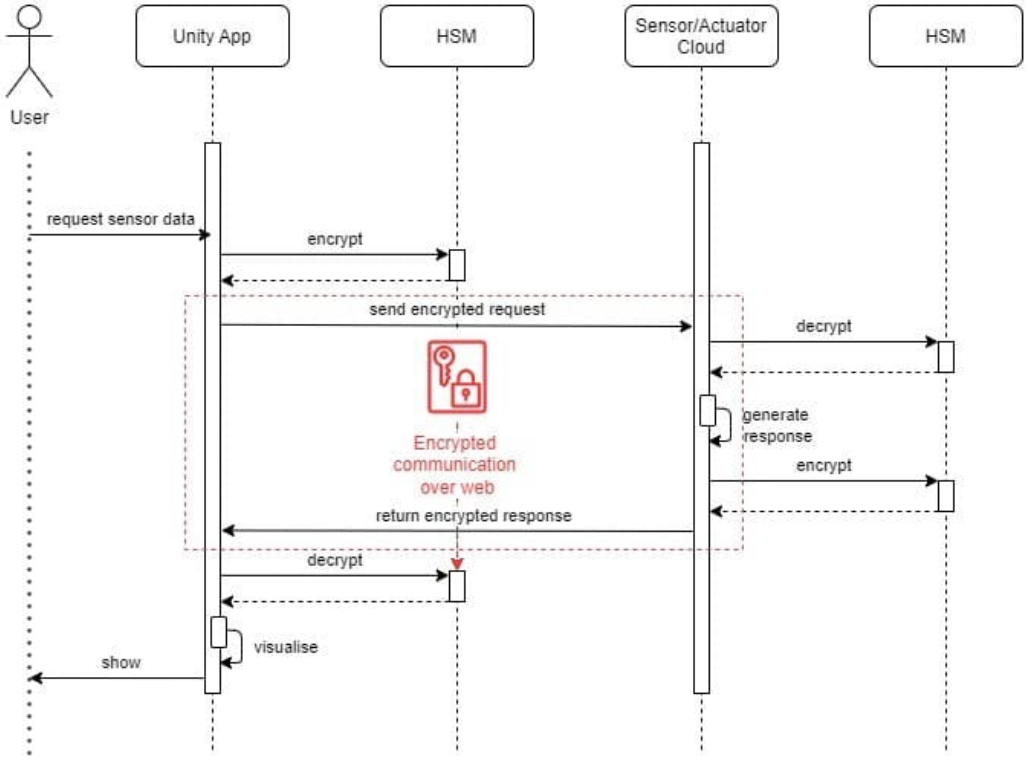

Figure 2. Encrypted data exchange between Client applications and Cloud

5.1.3 Threat Modelling of Physical Assets

TIoCPS employs enablers for physical resources/devices and gateways, and related trust & security means and platforms. The aim is to focus only on the physical resources/devices that are required in the context of the project UCs, and especially focus into the capabilities that are needed when enabling energy efficient and autonomous solutions for trust & security.

It is estimated that such solutions are needed in low power embedded sensors, actuators, and devices such as e.g., wearables/portables, aerial vehicles, robotic and other mobile asset devices and equipment's that shall operate as the key part of the CPS community's data spaces. The key operation in the CPS communities is related to the ownership, and therefore it is seen that authorization and policy services needed to take care of the access, communication, and information rights. The wireless asset device, gateways and the edge network side of the CPS system is especially challenging, because there is not necessarily always available connection to the infrastructures. These establish the other key challenges and responsibility area of the WP5 dealing with the trustworthy operations related to the data/information exposed from the physical world.

The physical CPS enablers are analysed in scope of Security Threat Modelling (STRIDE, see Table 3) and Privacy Threat Modelling (LINDDUN, see Table 4) methodologies to ensure trust and security in the TIoCPS platform.

Table 3. STRIDE threat model

Property	Description	STRIDE Threats
Authentication	The identity of users is established (or you are willing to accept anonymous users)	**S**poofing
Integrity	Data and system resources are only changed in appropriate ways by appropriate people.	**T**ampering
Non-repudiation	Users cannot perform an action and later deny performing it.	**R**epudiation
Confidentiality	Data is only available to the people intended to access it.	**I**nformation Disclosure
Availability	Systems are ready when needed and perform acceptably	**D**enial of Service
Authorisation	Users are explicitly allowed or denied access to resources.	**E**levation of privilege

Table 4. Properties used in privacy requirements classification associated with LINDDUN

Property	Description	LINDDUN Threats
Unlinkability	Hiding the link between two or more actions, identities, and piece of information (including cascade links)	**L**inkability
Anonymity	Hiding the link between an identity and an action or a piece of information	**I**dentifiability
Plausible Deniability	Ability to deny having performed an action that other parties can neither confirm nor contradict	**N**on-Repudiation
Undetectability and Unobservability	Hiding the user's activities	**D**etectability
Confidentiality	Hiding the data content or controlled release of data content	**D**isclosure of Information
Content Awareness	User's consciousness regarding his own data	**U**nawareness
Policy and Consent Compliance	Data controller to inform the data subject about the system's privacy policy, or allow the data subject to specify consents in compliance with legislation	**N**on-Compliance

5.1.4 Ontologies

The term ontology, which was borrowed from the domain of philosophy, describes the nature of being The ontology concept was initially utilized by artificial intelligence researchers who created simulations with the help of computers using automated reason-ing. From the IT perspective, ontology describes a set of depictive primitives in an explicit knowledge area Commonly adopted primitives include classes, qualities, and relationships together with their effects and boundaries. Ontology can be represented using a widerange of programming languages and schemes, such as description logic (DL), first-orderlogic, relational model, and UML The most commonly used ontology language in theinformation stack of the semantic web is web ontology language (OWL) [OWL originated from description logic to provide correct semantics of perceptions and associations.RDF ontology is also a part of the W3C technology information stack. It represents thesemantics of a domain in triple form (i.e., subject–predicate–object). The semantics denotedin RDF or OWL can be extracted by means of a query language known as SPARQL
https://www.researchgate.net/publication/366655724_Applications_of_Ontology_in_the_Internet_of_Things_A_Systematic_Analysis

5.1.5 Ontologies in IOT

This section presents several of the popular and most used ontologies in IoT. First, the semantic sensor network (SSN) ontology is based on the ontology design pattern, which describes the relations among sensors, stimuli, and observations. It also includes elementsfrom the stimulus sensor observation (SSO) design. SSN ontology can be understood fromfour major perspectives. (a) The sensor perspective describes what a sensor senses, howit senses, and what is being sensed. (b) The observation perspective refers to the dataunder observation and the associated metadata. (c) The system perspective discusses how asensor system is made up and deployed. (d) The feature and property perspective discusseswhat sensors sense about a property.Second, the suggested upper-merged ontology (SUMO) integrates a number of currentupper-level ontologies It contains different sections to describe ontology overall. Thefirst section is identified as the structural ontology that comprises descriptions of relationsthat help the framework describe the ontology properly. The second section is identifiedas the base ontology that includes essential ontological concepts, such as abstract objectsand the division between objects and methods. The set/class principle section of SUMO(i.e., the third section) contains simple set abstract knowledge. The numeric section describesbasic arithmetic tasks, and the temporal section builds relationships based on Allen's temporalrelations. The graph theory section offers general graph theoretic views. The unit of measuresection provides explanations of SI and other unit systems. The other sections of the ontologyprovide sub-hierarchies and axioms linked to process, object, and attribute types.Third, the sensor, observation, sample, and actuator (SOSA) ontology offers a lightweightcommon goal description of showing the collaboration among things as a part of perform-ing sampling, observation, and actuation SOSA is created by reusing the W3C SSNontology while considering user choice, specific audience, and technical requirements.SOSA can be used as an alternative to the SSN-based SSO principle. Fourth, the DogOntontology has particular importance in the interoperation between home automation sys-tems. The basic concepts of DogOnt are from real-world case studies and focus ondevice, state, and functionality modeling. This ontology can be described along five mainhierarchy trees: (a) building things, which demonstrates presented things that are eithermanageable or not; (b) building environment, which indicates where objects are situated;(c) state, which demonstrates the stable configurations that manageable things can ac-cept; (d) functionality, which reveals what manageable things can perform the action; and(e) home automation network component, which provides the specifications of each homeautomation network

https://www.researchgate.net/publication/366655724_Applications_of_Ontology_in_the_Internet_of_Things_A_Systematic_Analysis

5.2 Applications

5.2.1 HW Security

With the combination of hardware components; (i) Secure Gateway (SGW) and (ii) Hardware Security Module (HSM), the security of the shared data is enabled.
Secure Gateway (SGW) – SGW is a lightweight hardware device used on the Edge IoT side to ensure secure/encrypted data sharing from the edge device. Only the parties with eligible secret/private key can decrypt the data and use it (e.g. the central server with HSM).
Hardware Security Module (HSM) – HSM is a hardware device embedded to the central server via PCIe and provides cryptographic operations such as encryption/decryption, key generation, truly random number generation etc. It enables decryption of the messages received from the SGW (Figure 3), and it also provides a window for secure/encrypted data sharing with the third parties (Figure 4). Since not every partner/user can be equipped with the HSM device, a software solution, namely SoftHSM is provided to the users of the use-case.

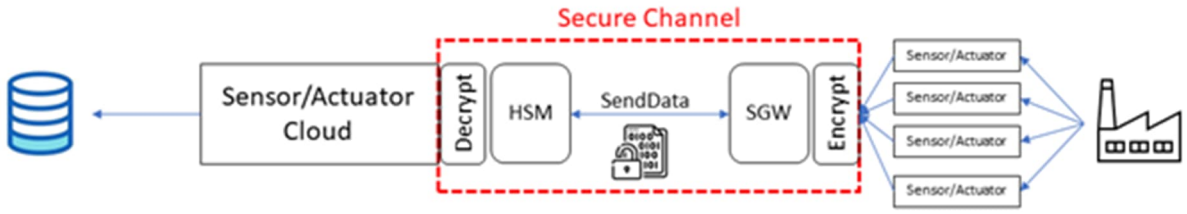

Figure 3. End-to-End Holistic Cyber-Physical Security.

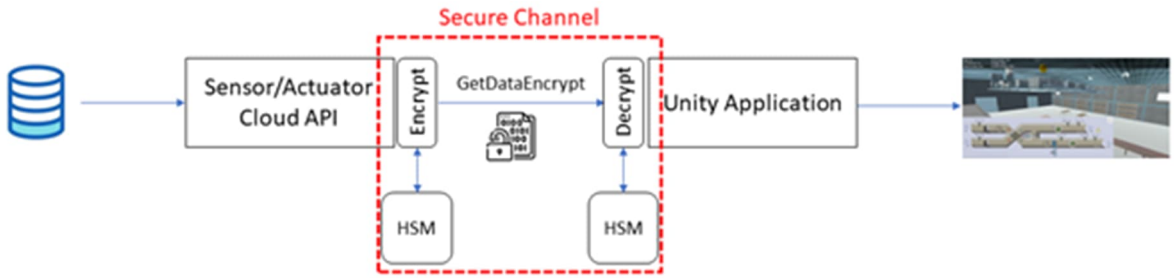

Figure 4. Encrypted data exchange between Client applications and Cloud.

5.2.2 AI Based Optimization

With the development of the industry, the electricity consumption of commercial buildings and the systems used in these buildings has gradually increased. Heating, ventilation and air conditioning systems, called HVAC, are controlled in various ways, allowing the temperature and ventilation of the environment to be adjusted. In this study, a fuzzy inference and machine learning based HVAC control system is proposed that is aware of the condition change and automatically adjusts the optimal conditions for the building occupants. The proposed system consists of two subsystems: ventilation and temperature control. The ventilation system uses a Random Forest algorithm that estimates the air quality index to provide fresh air. In the temperature control system,

Mamdani Fuzzy Inference System with four inputs and one output is used. Proposed architecture given in figure below.

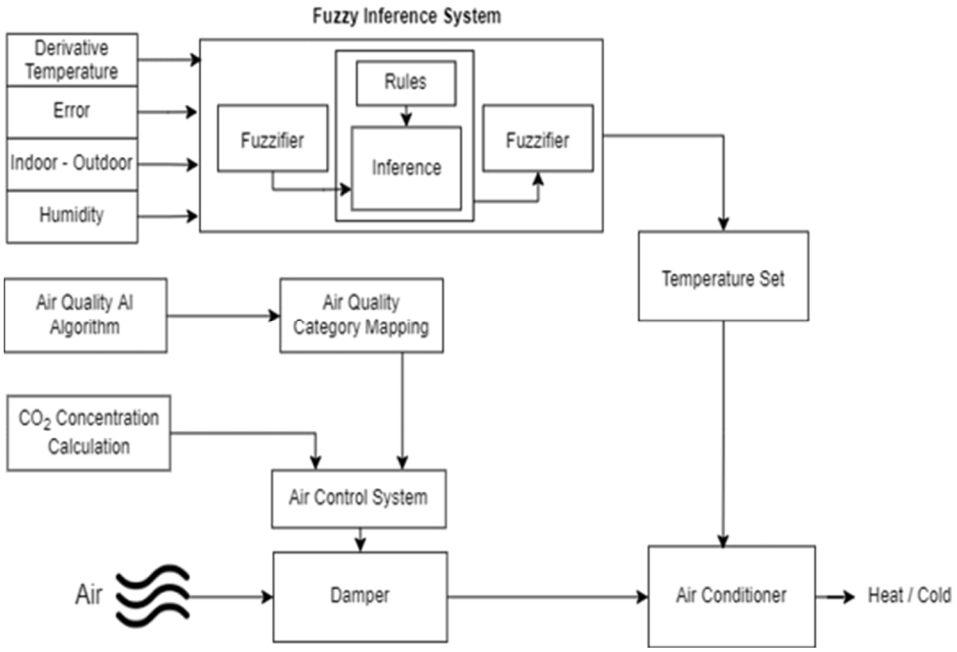

Figure 5. Proposed AI Based Architecture.

An edge device is a type of device located at the "end" of a network, usually close to the data source. Edge devices are used to collect data from sensors, cameras or other sources and then transmit that data to the cloud or a central location for processing and analysis. In this context, Jetson Nano was used as the end device. Jetson Nano is a small, low-power computer developed by NVIDIA. It is designed for use in embedded systems, robotics and other applications that require powerful processing capabilities in a compact form factor. Within the scope of the project, the ventilation system algorithms were installed in jetson nano in order to prevent delays and to make instant analyses at the points where intensive data transfer is required. Thus, early warning systems based on the Local Outlier Factor (LOF) technique can be used without excessive toxic gas accumulation in the environment and the user can be informed.

5.2.3 BIM Integration – Digital Twin

Instant observations of the HVAC sensors and actuators which were linked with the semantic backend (BIM + GIS + IoT ontology) could be monitored. The communication between the Sensor and Actuator Cloud and Digital Twin Application (Unity Application) was secure/encrypted via Hardware Security Module (HSM). The AI based optimization messages were sent to the Unity application directly.

For the final review, Air Quality sensors were made available for use on the Digital Twin application and they are added on the monitoring panel (see Figure 6).

An Event ontology is added on the semantic backend which supports the tracking of Anomaly (according to ENISA Thread Taxonomy) and Improvement (AI based optimisations) records (see Figure 7and Figure 8). The additional Event CRUD and query APIs support the event generation and client applications. The Digital Twin application has a polling rate of 20 seconds for checking whether an Anomaly or Improvement event record is registered in the Alp Aerospace Manufacturing Facility, through query APIs. As depicted

in Figure 9, as soon as an event record appears, it is rendered on the Digital Twin Application with the related information.

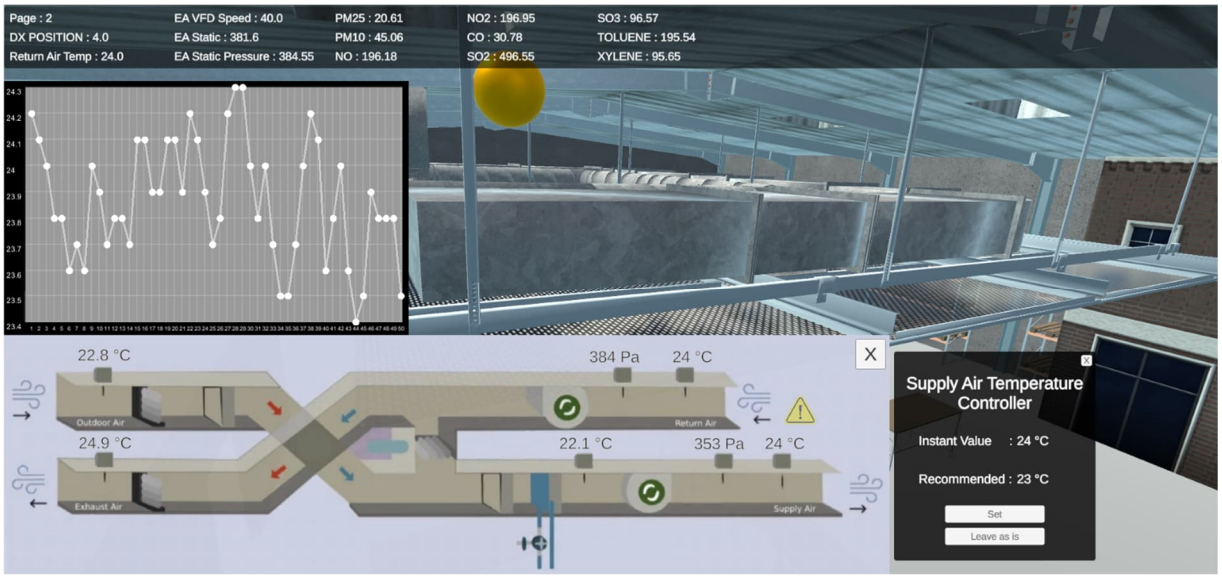

Figure 6. Unity-based Digital Twin Application with enhanced IoT data.

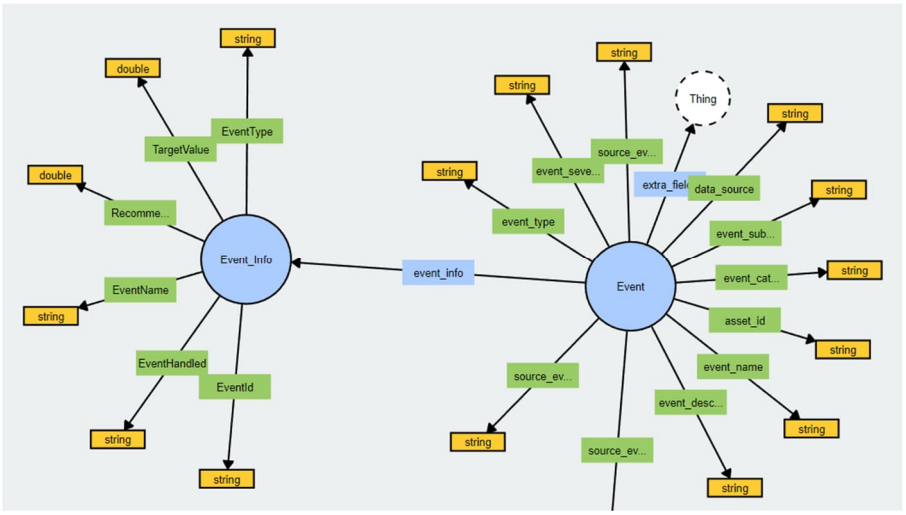

Figure 7. Event Ontology structure.

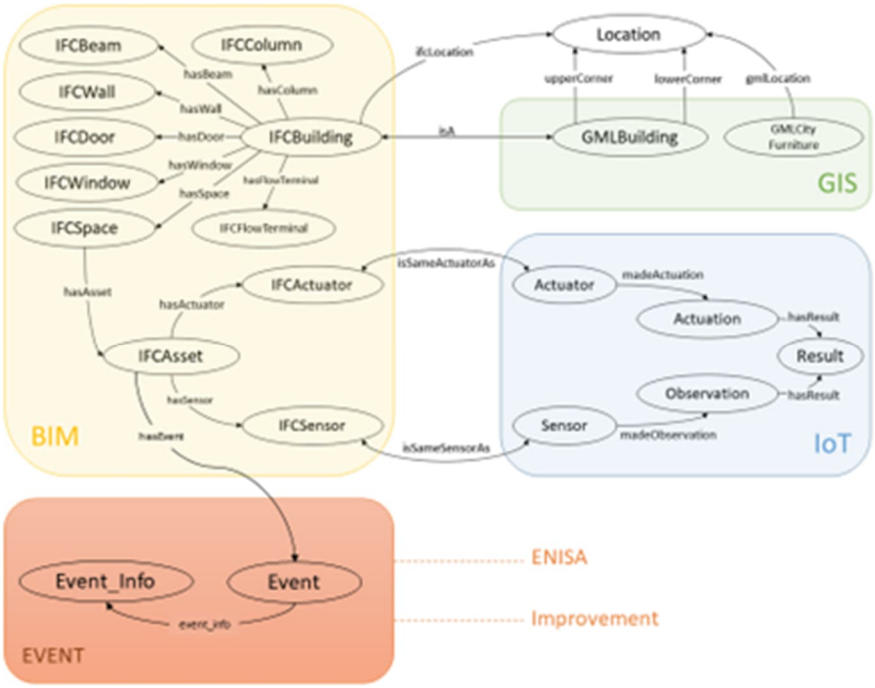

Figure 8. GIS-BIM-IoT-Event integrated super ontology.

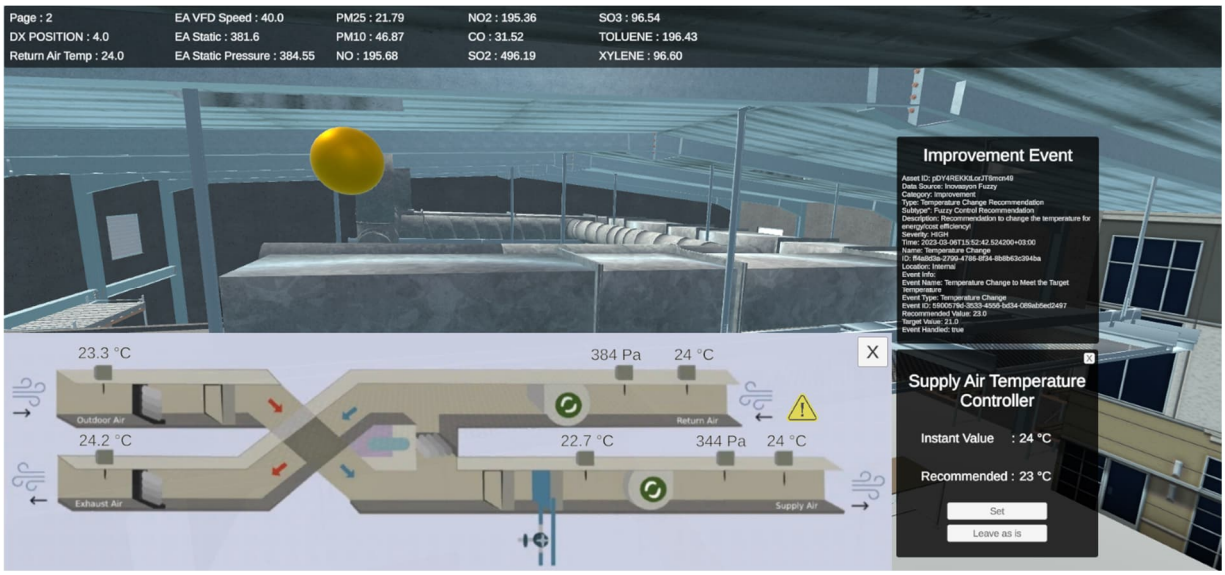

Figure 9. Event notification in the Unity-based Digital Twin application.

5.2.4 Live data visualization of IoT sensors using Augmented Reality (AR) and BIM

Building Information Modeling (BIM) is being widely used during the design and construction of buildings. Recently, the use of BIM technology during the operation and maintenance phase has been increasingly adopted. Additionally, Virtual Reality (VR) and Augmented Reality (AR) technologies are utilized in many applications during the design, construction, and the operation and maintenance phases. VR/AR visualizations of BIM models has been widely investigated, and their commercial applications are available in the market. On the other hand, there has been a surge of interest in the exploitation of the Internet of

Things (IoT) for constructed facilities, in the form of wireless networks connecting physical objects (e.g., sensing devices, facility assets, equipment, etc.). The generated data by IoT agents are aggregated, stored, processed and visualized for improving the safety and the utilization of facilities. Recent IoT systems offer data management and visualization solutions. However, to improve issues such as safety and indoor comfort conditions for facilities, most of the existing IoT deployment do not benefit from the enriched digital representations of BIM and its graphical visualization capabilities. Although the integration of sensor data with BIM models has been investigated in academia, storing such real-time data in a standard and structured manner remains to be further investigated. This research aims to visualize live environmental data collected by IoT agents in the AR environment built upon existing BIM models. In our case study, the environmental data, such as indoor air temperature, light intensity, and humidity are captured by sensors connected to Arduino microcontrollers. Sensor reading are then stored in the BIM model and visualized in both the BIM platform and the AR application in a real-time manner. The results of the case study showed that system can provide users with assistance for real-time monitoring of facilities' indoor thermal comfort condition.

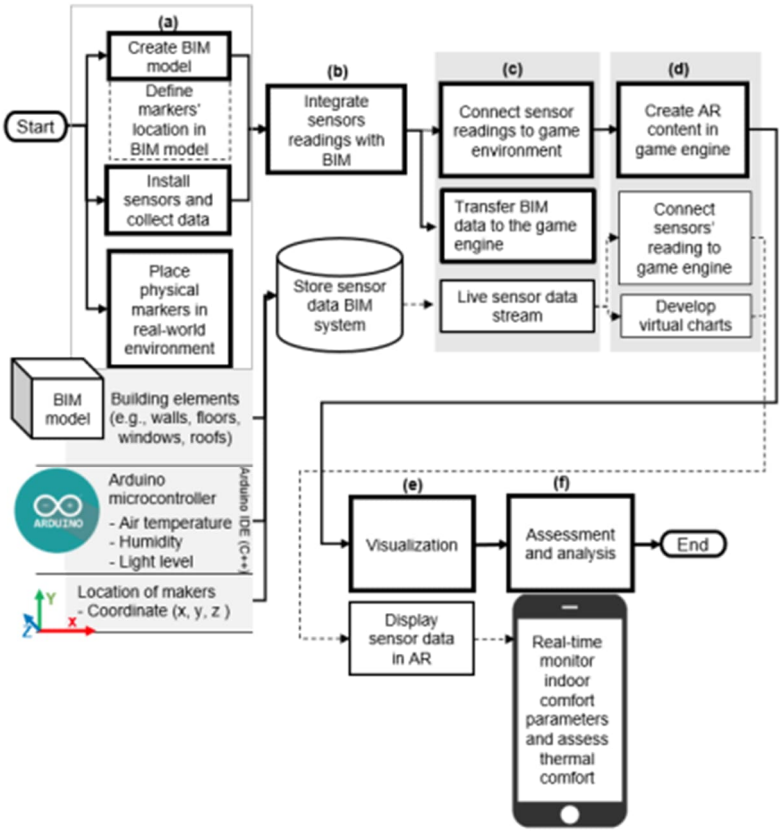

Figure 2. Process flow for integrating IoT sensors and BIM with AR

5.2.5 Interoperability between BIM and GIS

The main purpose of combining BIM with GIS is to integratestrong components of both systems for creating a data-rich virtualmodel of assets within the built environment (Song et al., 2017). BIMprovides detailed information about the geometry, materials, andother parameters of the asset. GIS provides spatial information thatcan be used for in-depth analysis of the built environment. BIM iswidely used for improving efficiency in design coordination andconstruction management. GIS has evolved into a robust system thatsupports spatial-temporal analysis and visualization of data sets forlarge-scale urban environments. Both BIM and GIS have tools andprocedures for visualizing data in 3D environment. ConsideringBIM can add detailed semantic information to GIS, and GISprovides geospatial context to BIM models, integrating BIM-GIShas great potential for facilities management by providing spatial-temporal analysis and management tools (Ohori et al., 2017; Songet al., 2017). BIM and GIS can be integrated at the data layer and the application layer. Data layer usually refers to the transfer ofgeometric and semantic information between the platform

5.3 Discussion

Digital twin is an evolving concept in the AEC industry. It providesa centralized platform for data collection, modification, analysis, andvisualization. Digital Twin is envisioned to have a geo-located and real-time weather data model, improving the accuracy in analysis comparedto the simulation-based analysis. To address the limitations in datacollection and analysis following the global pandemic, using IoT-basedsensors helps eliminating the need of an IEQ cart by facilitating spotmeasurements at specified intervals. Moreover, by setting up digitaltwins, researchers don't have to be physically present to takemeasurements. However, digital twins do not incorporate theaddition of human behavior to determine its impact on the buildingperformance and quantifying it would be challenging without thepresence of an observer (researcher) in the building.Integration of BIM and GIS technologies enhances the capabilitiesof the digital twin by adding comprehensive geometric, semantic,parametric, and built environment data. Moreover, high-levelapplications and use case scenarios have clarified the requirementsof a digital twin platform for POE. However, developing detailed usecase scenarios will help create a customized system UI catered to POE.The tests explain the step-by-step procedure followed for BIM-IOT-GIS integration
https://www.researchgate.net/publication/371522452_Applicability_of_BIM-IoT-GIS_integrated_digital_twins_for_post_occupancy_evaluations

5.4 References

1. Akanmu, A., Anumba, C., & Messner, J. (2013). Scenarios for cyber-physical systems integration inconstruction. Journal of Information Technology in Construction (ITcon), 18(12), pp. 240–260.Batty, M. (2018).
2. Digital twins. Environment and Planning B: Urban Analytics and City Science, 45(5), pp.Boje, C., Guerriero, A., Kubicki, S., & Rezgui, Y. (2020).
3. Towards a semantic Construction Digital Twin:Directions for future research. Automation in Construction, 114, 103179.Borrmann, A., Kœnig, M., Koch, C. & Beetz, J. (2018).
4. Building Information Modeling: Technology foundationsand industry practice, 1 ed. SpringerChang, K.-M., Dzeng, R.-J., & Wu, Y.-J. (2018).
5. An Automated IoT Visualization BIM Platform for DecisionSupport in Facilities Management. Applied Sciences, 8(7), 1086.Dave, B., Buda, A., Nurminen, A., & Främling, K. (2018).
6. A framework for integrating BIM and IoT throughopen standards. Automation in Construction, 95, 35–45.
7. Enabling Digital Twins with Advanced Visualization and Contextualization of Sensor Data with BIM and Web Technologies. Available from: https://www.researchgate.net/publication/354947245_Enabling_Digital_Twins_with_Advanced_Visualization_and_Contextualization_of_Sensor_Data_with_BIM_and_Web_Technologies
8. Richter, C. (2009). Visualizing Sensor Data. Trends in Information Visualization, 1–21. [2] Kim, T., Saket, B., Endert, A., & MacIntyre, B. (2017). VisAR: Bringing Interactivity to Static Data Visualizations through Augmented Reality, arXiv:1708.01377.

9. Latvakoski J., Heikkinen J. A Trustworthy communication hub for Cyber-physical systems. MDPI Future Internet Journal. *Future Internet* **2019**, *11*(10), 211; https://doi.org/10.3390/fi11100211 38p

6 Concluding Remarks

There are several challenges around the focused cyber-physical systems, but the lack of digital trust are estimated to be the most serious one. This is because it prevents the establishment of information sharing, and thus establishment of the data economy around CPS industries. This challenge is visible in several levels of the system, and thefore looking at the existing architecture system level technologies have been essential. In addition, the technologies available for the physical devices, communications and information level services have been essential. When thinking all of these levels, it is clear that the scope of the challenge field is very large. The approach for solving this has been limiting the scope by focusing into use cases from selected application sectors i.e. maintenance, energy flexibility, buildings, mobile and traffic services, and especially to the practical challenges around these cases. Therefore, we have focused this report to discuss only about the state of the art technologies, and state of the practise solutions avaible in the field related to these challenges and use cases. This means that the provided view to the state of the art and practises is just a snapshot of the available technologies related to cyber-physical systems.