Ref. Ares(2024)3931172 - 31/05/2024

EDIAQI

Evidence Driven Indoor
Air Quality Improvement

Deliverable D4.6

Privacy and IoT Security Version 1

Work Package 4

Pilots, data and campaigns

Version: Final

This project has received funding from the European Union's
Horizon Europe Framework Programme under grant agreement Nº 101057497.

Deliverable Overview

This deliverable collects the information and evaluates the IoT security as well as data privacy in the EDIAQI pilots. The deliverable assesses the IoT security measures to be applied in the EDIAQI pilots and makes recommendations on the IoT security measures to be applied. On the data privacy domain, the deliverable makes recommendations on how to ensure and improve privacy of the data collected in EDIAQI campaigns and pilots.

Additional Information

Type: R – Document, report

Dissemination Level: PU – Public

Official Submission Date: 31st of May 2024

Actual Submission Date: 31st of May 2024

Disclaimer

This document contains material, which is the copyright of certain EDIAQI partners, and may not be reproduced or copied without permission. All EDIAQI consortium partners have agreed to the full publication of this document if not declared “Confidential”. The commercial use of any information contained in this document may require a license from the proprietor of that information. The reproduction of this document or of parts of it requires an agreement with the proprietor of that information.

This project has received funding from the European Union’s
Horizon Europe Framework Programme under grant agreement Nº 101057497.

Document Revision History

Version	Date	Description	Partners
V0.1	17th of February 2024	First draft	THIN
V0.2	29th of February 2024	Document revised	THIN
V0.3	15th of March 2024	Document revised	THIN
V0.4	29th of March 2024	Document revised	THIN
V0.5	5th of April 2024	Document revised	THIN
V0.6	12th of April 2024	Document revised	THIN
V0.7	26th of April 2024	WINGS solution description improved FROST server description added	LAS WINGS
V0.8	6th of May 2024	Document revised, future plans added	THIN
V0.9	16th of May 2024	Document revised	THIN
V1.0	28th of May 2024	Document revised	KNOW
Final	31st of May 2024	Quality check	LC

Authors and Reviewers

Authors and contributors

- Alessandro Battaglia (LAS)
- Gianna Karanasiou (WINGS)
- Jürgo Preden (THIN)
- Heikki Kitt (THIN)
- Kuldar Loime (THIN)

Reviewers

- IMROH
- NIB

This project has received funding from the European Union's
Horizon Europe Framework Programme under grant agreement Nº 101057497.

Statement of Originality

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation, or both.

This project has received funding from the European Union's
Horizon Europe Framework Programme under grant agreement № 101057497.

Table of Contents

This project has received funding from the European Union's
Horizon Europe Framework Programme under grant agreement Nº 101057497.

This project has received funding from the European Union's
Horizon Europe Framework Programme under grant agreement Nº 101057497.

List of Figures

List of Tables

This project has received funding from the European Union’s
Horizon Europe Framework Programme under grant agreement Nº 101057497.

List of Terms and Abbreviations

Abbreviation	Description
6LowPAN	IPv6 over Low-Power Wireless Personal Area Networks
ANT	Institute of Anthropological Research
API	Application Protocol Interface
AQ	Air Quality
CAS	Chemical Abstracts Service
CoAP	Constrained Application Protocol
CEN	Comité Européen de Normalisation
CQL	Common Query Language
CRS	Coordinate Reference System
CSV	Comma Separated Values
EC	European Commission
ECHA	European Chemical Agency
EIONET	European Environment Information and Observation Network
FROST	FRaunhofer Opensource SensorThings
FTP	File Transfer Protocol
GEMET	GEneral Multilingual Environmental Thesaurus
GML	Geography Markup Language
HTTP(S)	Hypertext Transfer Protocol (Secure)
INSPIRE	Infrastructure for Spatial Information in Europe
IAQ	Indoor Air Quality
IOT	Internet Of Things
IPChem	Information Platform for Chemical Monitoring
ISO	International Organization for Standardization
JPEG	Joint Photographic Experts Group
JSON	JavaScript Object Notation
KML	Keyhole Markup Language
KVP	Key Value Pair

LOD	Level Of Details
MQTT	MQ Telemetry Transport
OGC	Open Geospatial Consortium
OWS	OGC Web Services
PDF	Portable
PNG	Portable Document Format
REST	Representational state transfer
RDBMS	Relational Database Management System
SensorML	Sensor Markup Language
SLD	Styled Layer Descriptor
SQL	Structured Query Language
SSN	Semantic Sensor Network
STA	SensorThings API
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
W3C	World Wide Web Consortium
WGS84	World Geodetic System 1984
WMS	Web Map Service
WFS	Web Feature Service
XML	eXtensible Markup Language

1. Executive Summary

This deliverable introduces the importance on IoT security and privacy issues, including the context of indoor air quality data collection. The deliverable discusses data privacy and security issues for IoT systems, such as indoor air quality data collection systems that are applied in the EDIAQI project.

The data privacy recommendations discussed in the deliverable follow privacy and security guidelines generally accepted by the industry and adopted by market stakeholders.

Methods for data anonymization are discussed to provide input to the EDIAQI project for anonymization of data collected on indoor climate quality.

The deliverable describes and assesses the security measures applied for the IoT systems in the pilots and campaigns in EDIAQI and makes preliminary recommendations on modifying the security measures. The security measures applied are discussed on a per-campaign basis, describing and evaluating the measures applied in each technical solution applied in the EDIAQI project. The security and privacy features for each technical solution are discussed in detail.

Additionally, the deliverable makes recommendations on data privacy for the sensor data collected in the EDIAQI pilots and campaigns and delivered to the central data platform.

The deliverable provides an overview of the plans for added topics for the next EDIAQI solution security deliverable.

2. Introduction

2.1 Importance of data security and privacy in IoT

Data security is important in Internet of Things (IoT) for a range of reasons, which are relevant in ensuring data security and privacy in the case of the EDIAQI project pilots. In this section the importance of data security is discussed to provide the motivation for the related work.

IoT devices can also collect a vast array of data from users, ranging from personal preferences, environmental data on buildings and environments, to detailed information about one's physical activities and locations. This data collected by IoT devices can contain sensitive information about the users, as well as the monitored environments that the users may not want to share with third parties. In the EDIAQI project, the indoor climate data most certainly can reveal private data about the individuals, their life and habits. It is therefore essential to protect this data from being accessed by third parties and to ensure the privacy of the data is preserved when the data is being processed for research purposes. Preserving privacy of IoT data is important for maintaining trust and acceptance of the systems, which is important to encourage the adoption of IoT systems. Users are more likely to embrace and use IoT devices if they believe their personal information is handled securely and with respect for their privacy. They are also more likely to adopt the solutions which, in the case of EDIAQI, helps to protect the health of the individuals. Potential privacy breaches can undermine this trust, potentially leading to resistance against adopting new technologies.

Naturally, the legal and ethical obligations for data privacy must be followed in any case, as the GDPR legislation states clearly how the private information should be handled. Any solutions must comply with this legislation to avoid legal consequences and to maintain ethical standards in the operations. Therefore, privacy protection is not just a technical requirement, but a legal and moral one as well.

Protecting privacy in IoT also empowers users by giving them control over their own data. This includes deciding what information to share, with whom, and in what context. Privacy protection supports the autonomy of individuals and allows them to engage with IoT technology on their own terms. Data privacy in IoT ensures the security and dignity of individuals while encouraging a trustworthy and legally compliant environment for technological advancement.

2.2 Rationale

The deliverable D4.6 reports on the activities of Task T4.6 on security and privacy. The motivation for ensuring the security and privacy was outlined in the previous section. The deliverable presents the sensor data security and privacy aspects, reviews the security methods applied and provides inputs to the project partners on ensuring the privacy of the collected data.

2.3 Deliverable objectives

The objective of the deliverable is to provide an overview of data security and privacy and to provide recommendations on security and provide recommendations for ensuring data privacy of the data collected in EDIAQI pilots and campaigns. The deliverable helps to guarantee that data collected in EDIAQI pilots can be shared and made accessible without any privacy violations.

3. Data privacy

3.1 Importance of data privacy

Sensing devices, such as IoT devices (including devices monitoring indoor climate quality) can gather detailed information about an individual's daily habits and activities, including consumption rates, location, and health-related data. Similarly, sensing devices can gather detailed information on the routines of organisations, which can be used by to influence the operation or competitiveness of an organisation or plan malicious activities, such as terror attacks.

The indoor climate sensing devices employed in the EDIAQI project collect information about indoor climate quality, while it may seem that these data are not personalised and that the privacy of these data is not critical, it is not the case. The indoor climate data can be also used to deduce information about the daily habits and activities of individuals and organisations. Information extracted from analysing the data on routines of individuals and organisations could be used maliciously, therefore adversely affecting the individuals and organisations.

It is important to ensure data privacy to fulfil legal and regulatory compliance requirements – in the EU the GDPR regulations require that private data is not shared without the data owner's consent.

When data privacy is maintained, trust is built between users and IoT service providers, which encourages adoption and continued use of IoT devices and systems by users. This helps to deliver the value of the IoT solutions to the users. Conversely, data breaches or privacy violations can severely damage the reputation of IoT solutions and deter potential users. Therefore, in order to support the use of IoT indoor climate sensors for identifying health threats and to encourage citizens to be better informed and to demand healthier environments, the IoT solutions for indoor air quality monitoring must adhere to data privacy principles. Also, the information on security and privacy must delivered to the users – the users must feel that their data is processed securely and that their privacy is preserved.

Potentially ignoring data privacy enables IoT solution providers to explore opportunities for data monetisation without user consent. Private data on individuals and organisations could be also utilised for what is called surveillance capitalism – companies planning their sales, marketing and other activities by utilising data that should remain private. If left unchecked, the increasing volume and variety of IoT data in conjunction with traditional and nascent business models, may lead to an expansion of surveillance capitalism with even more far-reaching consequences. This threat has also been identified by UNESCO Data privacy and the Internet of Things | UNESCO Inclusive Policy Labi.[1] When organisations demonstrate a commitment to data privacy, they can attract potential collaborators and unlock new possibilities while ensuring the privacy rights of their users.

Ignoring data privacy also potentially enables individuals and organisations with criminal intent to make use of the private data for conducting criminal activities. A range of criminal activities could be planned to utilize the information extracted from the private IoT data, so ensuring security and preserving privacy is important also from this perspective.

Respecting data privacy is an ethical responsibility for solution providers that rely on IoT technology. It reflects a commitment to protecting the rights and dignity of individuals and upholding ethical principles in data collection, storage, usage and visualisation.

Any IoT solution providers must provide users the ability to control how their IoT data is shared and determine who can access it. In the IoT domain, privacy carries strong implications of trust, transparency, and control:

- The individuals should be able to control how the information collected by the IoT devices is shared and determine who has access to the data from devices in the homes, workplaces, schools and kindergartens, in your car, and on your person.
- There should be clarity about how information about people is collected, used, and shared with others. IoT devices and their applications should enable the user to find out what information is collected and shared, when and with whom.

[1] Data privacy and the Internet of Things | UNESCO Inclusive Policy Lab

This project has received funding from the European Union's
Horizon Europe Framework Programme under grant agreement № 101057497.

- The IoT solutions that provide access to data to third-party individuals and organizations must ensure data pseudonymisation or anonymisation when data is shared with third parties. As this is the case with EDIAQI, anonymisation must be ensured.

The Internet Society has published an IoT policy brief for policymakers[2], which serves as a good input for shaping the policy for IoT privacy and security. Safeguarding data privacy ensures trust, protects individuals' rights, and prevents misuse of sensitive information in the context of IoT systems collecting indoor climate data.

3.2 Data Privacy Principles

Ensuring data privacy is a complex task that must be tackled comprehensively across different system design stages, domains and system components. Privacy should be achieved by design by integrating privacy preserving principles into the design of IoT systems from the outset. So, the privacy implications should be considered during system architecture, development, implementation as well as operation.
Balancing innovation with privacy is essential for successful IoT and Edge system deployments and utilization. By safeguarding data privacy, smarter and more respectful built environments can be created, which are also welcomed by the users.
Below the principles for data privacy are discussed in more detail.

3.2.1 Data minimisation

Starting from the system design and planning stages: an important aspect of data privacy is data minimisation - only necessary data should be collected that is needed to achieve the desired outcomes. So, excessive data collection should be avoided, which includes the temporal granularity of data, modalities of data and locations from where data is collected.

- Optimal temporal granularity – data collection intervals should be optimised to obtain a sufficient amount of data points required for characterising the phenomena

[2] Policy Brief: IoT Privacy for Policymakers - Internet Society

This project has received funding from the European Union's
Horizon Europe Framework Programme under grant agreement Nº 101057497.

being monitored. If data is collected more frequently, the granularity is too high, which means that data is collected excessively.

- Optimal data modalities – only the data modalities (or data types) required for characterising the observed physical phenomena should be collected. This ensures that phenomena, which was not supposed to be monitored is not monitored.
- Optimal locations – data should be collected only data from the locations that characterise the monitored environment or phenomena. This means that sensors should be placed only to such locations and the sensors should only observe these areas.

While in the context of exploring new domains in research sometimes the approach of "collecting as much data as possible" is applied, this principle is not applicable in case of IoT systems deployed with users as in this case the fundamental research that provides responses to fundamental questions on which data characterises which phenomena should already have been answered. The *data to decision* principle from the domain of situation management should be followed – only data needed for achieving specific decisions or analysis results should be collected and stored.

While supporting data privacy, the data minimisation principles also support cost effectiveness and optimal system design.

3.2.2 Security of data transmissions and data management

A large part of data privacy is also security of data transmissions and management. Data should be encrypted during transmission between sensors, gateways, and cloud servers. Similarly, data should be secured also when it is stored in locations and the access to data should be protected by established authorisation and authentication methods. The encryption, authentication and authorisation methods applied should follow the established standards and best practices to ensure that the security of data is guaranteed.

Part of ensuring security of data transmissions and data management is up-to-date software and firmware, so firmware (and its components) in the devices and the software (and its components) in the servers must be updated regularly to address any possible security vulnerabilities. The security of data transmissions falls under the IoT security domain, which is addressed in section 4.2 of this deliverable in detail.

3.2.3 Privacy through anonymisation

Once the data security is ensured in the communications and data management, the privacy if the users and organisations should be addressed to ensure that the authorised users of the systems do not have access to private data of individuals and organisations. The anonymisation methods applied should minimise the possibility of re-identification of individuals and organisations based on the collected and stored data. The applied anonymisation methods should ensure that the individuals and organisations within a group cannot be distinguished based on the collected information. Identification could be achieved by analysing the collected data values if these can be attributed to specific individuals. However, this is not likely in the case of indoor climate data as similar indoor climate conditions could occur in multiple buildings and rooms.

3.2.4 User consent and transparency

One of the important principles in data privacy is consent and transparency, which must be maintained when handling the user data. For any user data consent must be obtained from individuals before the data can be used for any purposes. The consent obtained from users must contain an explanation on how and for what purposes their data will be used. Additionally, the data collection practices, purposes, and retention policies must be transparently communicated to users at the level understandable by the users. This ensures that the users are aware of how their data is handled and that the consents provided are informed.

4.2 Methods for data security

IoT data security must be addressed at all layers of the computing stack starting from the physical world where the devices are to all the way to the Cloud where the data is stored, processed, shared and visualised.

4.2.1 Device level data security

At the device level the hardware and local communications must be designed in a secure way, ensuring that eavesdropping or data injection and other attacks could not be implemented in an easy way. The Edge processing in the devices must be also implemented in a secure way, ensuring that data breaches from this origin would not be possible. Any

This project has received funding from the European Union’s
Horizon Europe Framework Programme under grant agreement № 101057497.

communication from the devices to a gateway or a data collection point (which in case of WiFi communications would be the network router) should be secure. The security methods applied should eliminate the possibility of many types of attacks, such as packet sniffing, rogue access points, packet injection, man in the middle, spoofing or encryption cracking. Naturally, not all types of attacks can be eliminated, like Denial of Service or jamming, but these are not critical from the perspective of privacy.

Most privacy threatening attacks can be eliminated when state of the art security measures have been implemented and correctly applied. Naturally, the solution operator must also ensure that the device software is up to date so that any detected security vulnerabilities would not be exploitable. Preferably secure remote updating of software should be implemented, which ensures that the device software can be updated when the need arises.

4.2.2 Gateway level data security

At the gateway level the security principles applied are established methods that are also used in mainstream devices, such as network routers. The gateway devices should initiate the upstream connections to the servers and any management interfaces available at the gateway devices should be secured using industry best practices. Any certificates or other security credentials should be stored in the gateways in a secure way, minimising the possibility of exploitations.

The messaging to and from the gateways should be managed, reducing the possibility of overloading the network and ensuring controlled delivery. Also, in case of the gateways the solution operator must also ensure that the gateway software is up to date so that any detected security vulnerabilities would not be exploitable. Preferably secure remote updating of gateway software should be implemented, which ensures that the device software can be updated when the need arises.

4.2.3 Cloud level data security

At the Cloud level, industry best practices for securing Cloud solutions should be applied. The identification, authentication and authorisation methods applied should ensure that only authorised individuals and technical systems have access to the Cloud system. Any user or API access should be encrypted to ensure that data is not compromised during transmissions.

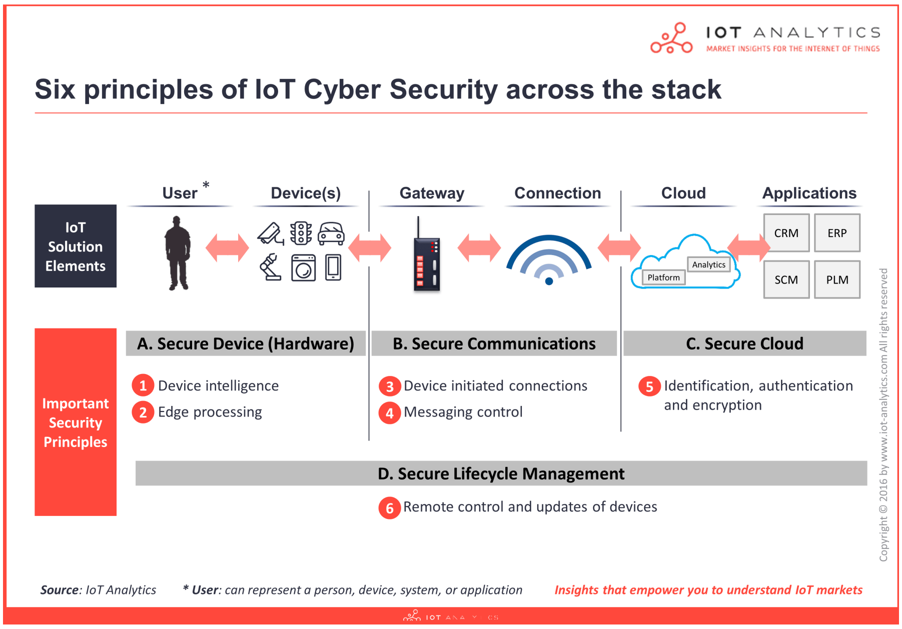
IOT ANALYTICS
MARKET INSIGHTS FOR THE INTERNET OF THINGS
Six principles of IoT Cyber Security across the stack
IoT Solution Elements
User *
Device(s)
Gateway
Connection
Cloud
Applications
Analytics
Platform
CRM
ERP
SCM
PLM
Important Security Principles
A. Secure Device (Hardware)
1 Device intelligence
2 Edge processing
B. Secure Communications
3 Device initiated connections
4 Messaging control
C. Secure Cloud
5 Identification, authentication and encryption
D. Secure Lifecycle Management
6 Remote control and updates of devices
Source: IoT Analytics
* User: can represent a person, device, system, or application
Insights that empower you to understand IoT markets
Copyright © 2016 by www.iot-analytics.com All rights reserved

Figure 1 – Six principles of IoT security (source: IoT Analytics)

4.3 Methods for data anonymisation

Being an important component of data privacy, the preservation of anonymity has a high importance in the context of indoor climate data collected by IoT sensors. So strong data anonymity methods must be applied when data on indoor climate data is shared with third parties or when these data are used for research.

In the data privacy domain, the concept of K-anonymity is used. By applying k-anonymity, the risk of re-identification of individuals and organisations based on the data is reduced because individuals and organisations within a group cannot be distinguished based on the available information. However, it's important to note that k-anonymity is not a perfect solution and has limitations. For example, it does not protect against certain attacks like attribute disclosure or background knowledge attacks. It must be also noted that improper implementation or inadequate choice of k value can compromise privacy or utility.

This project has received funding from the European Union's
Horizon Europe Framework Programme under grant agreement Nº 101057497.

Therefore, k-anonymity should be used as part of a comprehensive privacy protection strategy alongside other techniques and safeguards.

A dataset is considered k-anonymous if quasi-identifiers for each entity (individual or organisation) in the dataset are identical to at least k – 1 other people also in the dataset[3].

In a dataset, each record typically consists of several attributes that characterise the data record. K-anonymity is achieved by generalizing or suppressing certain attributes in a way that groups of records are indistinguishable from each other, while still providing meaningful data for scientific analysis or analytics targeted for end-users.

K-anonymity is applied using the following steps:

- Group Formation: Records in the dataset are grouped together based on their shared attributes. Each group should contain at least k records.
  In the context of EDIAQI, data the dataset groups are data collected from specific locations.
- Attribute Generalisation: Within each group, certain attributes are generalised or modified to make them less specific. As an example, numerical values might be replaced with ranges, and categorical values might be replaced with more general categories.
  In the context of EDIAQI data the collected sensor data should not be generalised but the addresses of the locations should be generalised to preserve anonymity
- Sensitivity Reduction: Sensitive attributes that could potentially identify individuals are either generalized or suppressed entirely. This ensures that each group of records remains indistinguishable from others.
  In the context of EDIAQI data the physical addresses of the data collection points can be identified as sensitive attributes.

[3] Computing k-anonymity for a dataset: https://cloud.google.com/sensitive-data-protection/docs/compute-k-anonymity

- Preservation of Utility: While ensuring privacy, k-anonymity aims to preserve the utility of the dataset for analysis. This means that the anonymised data should still be useful for legitimate purposes, such as statistical analysis or data mining.
  In the context of EDIAQI data the utility can be preserved by not generalising the collected sensor data.

This project has received funding from the European Union's
Horizon Europe Framework Programme under grant agreement № 101057497.

5. Data Privacy in EDIAQI

This chapter describes the data privacy and security topics as they are applied in the EDIAQI project. The data privacy aspects of data minimisation are addressed in EDIAQI by:

1) Collecting only essential data on indoor air quality parameters.
2) Collecting data only at the granularity required to provide a comprehensive overview of the indoor climate quality.
3) Placing sensors only in areas that need to be monitored in the EDIAQI pilots.

Data anonymisation is addressed in EDIAQI at the Cloud level to ensure that all the data collected could not be traced back to the individuals, buildings and organisations from where the data was collected.

The main focus of this chapter is on data security as ensuring that security is one of the basic building blocks of data privacy and it provides additional privacy in additional to the data minimisation. The chapter contains the description of IoT security measures applied in individual pilots and evaluates the security.

5.1 Security aspects assessed in pilots

In assessing the security, Table 1 presents the aspects are assessed.

Table 1 IoT Security aspects addressed in EDIAQI

System level	Devices	Important aspects
Cloud security	Cloud servers	Industry best practices applied for data security Identification, authentication and authorization methods applied Any communicated data encrypted
Communication from the Edge to the Cloud	Gateways routers	Upstream connections initiated by gateways, authenticated Messaging managed and encrypted Software updateable remotely via a secure channel
Communication from sensors	IoT sensors	Secure local communication, main attack vectors eliminated Software updateable securely

5.1.1 Measurement and monitoring implemented by Lab Service Analytics

5.1.1.1 General communication architecture design

EDIAQI Project pilots employing Lab Service Analytica's NetPID™ technology rely on the use of sensors that send authenticated and authorised data through services offered by Amazon AWS.

Data read from the sensors are uploaded to the cloud every two minutes, and the sensors can receive commands from the cloud to change configurations. Security measures are in place to ensure that only authorised data can be accessed. For this, a certificate must be entered to enable authentication and authorisation management for cloud access through an application gateway. This service uses support provided by AWS, called "Cognito," to ensure that data is accessed securely.

5.1.1.2 Data communication from sensor to gateway

Each NetPID™ remote system can connect to internet through a secured Wi-Fi network (WPA2 security). Credentials for accessing the network are stored in the memory of each instrument in an encrypted manner and consequently cannot be used by external agents. Data communication is via the HTTPS protocol, which uses TLS encryption to ensure authentication, data integrity, and confidentiality. This protocol allows secure communication from source to recipient over TCP/IP networks, such as internet.

When accessing the cloud, generated certificates are used to negotiate session keys for individual devices or users. This provides a high level of security because any authorised remote device can only access its own data. This prevents the possibility of sensitive data being intercepted by unauthorised third parties.

5.1.1.3 Data communication within the Cloud

Data communication within the cloud is fully managed using AWS services. Data exchange between the various services takes place within Amazon's network architecture, which ensures a high level of data security. Regarding the security of the AWS cloud, it should be noted that the company takes several security measures to ensure data protection. In particular, the data centres used for data storage comply with the highest security standards, such as ISO 27001, SOC1 and SOC2. In addition, AWS provides several advanced security tools, such as encryption of data in transit and at rest, multifactor authentication, and granular access control. Finally, AWS ensures the reliability and business continuity of the service through the presence of data backup and recovery systems, and a global distribution network that ensures 24-hour service availability.

5.1.1.4 Data communication from the Cloud to the user interface or external systems

Data stored in the Lab Service Analytica cloud are accessible to the user via a web-accessible dashboard. The interface uses software libraries (SDKs) provided by AWS to enable secure mode access to the cloud. Communication is via HTTPS protocol that ensures data security. Another way to access the data is by using the AWS API Gateway service, which is used by third-party organisations to connect to the cloud without going through the methodologies described above. An IAM user is generated with which the organisation can make secure calls to the cloud.

In light of the needs to apply innovative solutions for data interoperability to the project, Lab Service Analytica together with and in accordance with the evaluations made by Deda Next (Lead Partner of the activity), will develop one of two solutions (one does not exclude the other) for data transfer via the Fraunhofer Institute's SensorThingsAPI standard (recommended by the European Commission as "good practice" for distribution of measurement data according to INSPIRE guidelines).

The use of open-source technology makes the road to innovation easier. The FROST®-Server ("Fraunhofer Open Source SensorThings API Server") has high performance with low resource consumption and an openness that facilitates use in both the research environment and commercial applications.

5.1.2 Measurement and monitoring implemented by WINGS

5.1.2.1 General communication architecture design

The communication architecture and the data security on WINGS systems are described in detail on three different levels (device, communication, and protocols) and the communication architecture and data security of AIRWINGS is summarised in **Error! Reference source not found.**.

This project has received funding from the European Union's
Horizon Europe Framework Programme under grant agreement Nº 101057497.

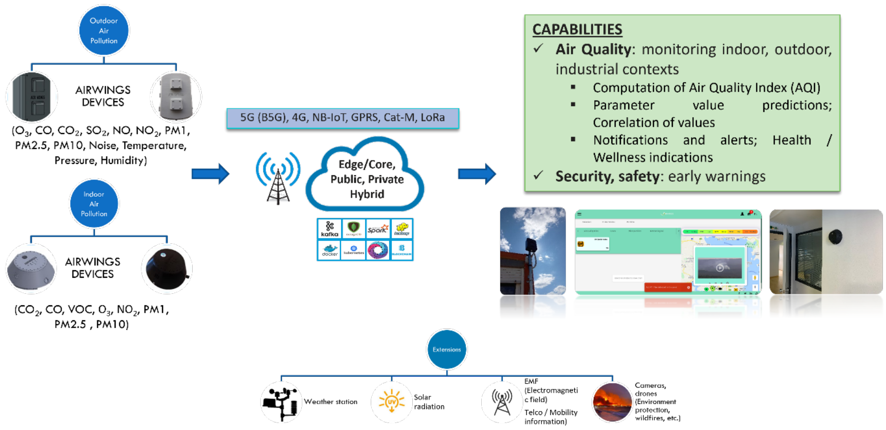
Outdoor Air Pollution
AIRWINGS DEVICES
(O_3, CO, CO_2, SO_2, NO, NO_2, PM1, PM2.5, PM10, Noise, Temperature, Pressure, Humidity)
Indoor Air Pollution
AIRWINGS DEVICES
(CO_2, CO, VOC, O_3, NO_2, PM1, PM2.5 , PM10)
5G (B5G), 4G, NB-IoT, GPRS, Cat-M, LoRa
Edge/Core, Public, Private Hybrid
kafka
Spark
docker
kubernetes
BLOCKCHAIN
CAPABILITIES
✓ Air Quality: monitoring indoor, outdoor, industrial contexts
▪ Computation of Air Quality Index (AQI)
▪ Parameter value predictions; Correlation of values
▪ Notifications and alerts; Health / Wellness indications
✓ Security, safety: early warnings
Extensions
Weather station
Solar radiation
UV
EMF (Electromagnetic field)
Telco / Mobility information)
Cameras, drones (Environment protection, wildfires, etc.)

Figure 2. The communication architecture and data security of AIRWINGS

5.1.2.2 Device level

At the device level, it is possible to implement the AES (Advanced Encryption Standard) encryption protocol. AES stands as one of the most robust encryption standards available and serves to safeguard the data stored on the device. The process of data encryption serves as a fundamental security measure, guaranteeing that even if an unauthorised individual gains physical access to the device, the confidentiality and integrity of the stored data remain uncompromised.

5.1.2.3 Communication level

At the communication level, the security of the device relies on the security provided by the narrowband network. The narrowband network operates within an exceptionally limited frequency bandwidth, thereby substantially impeding the ability of malicious entities to eavesdrop on or intercept the transmitted data. This aspect contributes an additional stratum of security that fortifies the device's communication capabilities.

WINGS systems also offer support for Wi-Fi connectivity, providing a versatile option for data transmission. Utilising Wi-Fi, devices within the network can seamlessly communicate over local area networks (LANs) or connect to the internet, facilitating rapid and efficient data exchange. To ensure the security of Wi-Fi communication, robust encryption standards such as WPA2 (Wi-Fi Protected Access 2) are implemented, safeguarding data integrity and confidentiality against potential threats.

Furthermore, devices are designed to accommodate diverse networking environments, including 3G, 4G, and even the latest 5G cellular networks, as well as LoRa (Long Range) and GPRS (General Packet Radio Service) technologies. This broad compatibility ensures that WINGS devices can seamlessly integrate into a wide array of infrastructures, enabling connectivity across different geographical regions and varying network capabilities. Whether operating in urban environments with high-speed cellular networks or remote areas with low-power, long-range connectivity requirements, WINGS systems adapt to the specific networking needs, ensuring reliable communication and data exchange.

5.1.2.4 Protocol level

At the protocol level, the use of secure UDP (User Datagram Protocol) helps maintain the integrity of the communication. Every data packet dispatched by the device is accompanied by a corresponding acknowledgment, serving as a definitive confirmation of accurate transmission. This process ensures that data remains uncorrupted and imparts a mechanism through which the devices can stay apprised of the status of their transmitted data, thus eliminating the potential for information loss.

In summary, the communication architecture and data security of WINGS systems are designed with multiple layers of protection. At the device level, AES encryption protocol ensures robust security for stored data. The communication level encompasses various options, including the security provided by narrowband and Wi-Fi connectivity, which utilizes encryption standards like WPA2 for secure data transmission. Moreover, WINGS systems support a wide range of networks, including 3G, 4G, and 5G cellular networks, as well as LoRa and GPRS technologies, ensuring adaptability to diverse networking environments. Additionally, at the protocol level, secure UDP helps maintain data integrity during transmission, further enhancing the overall security posture of the device. This comprehensive and layered approach ensures the secure operation of WINGS systems across different communication scenarios.

Table 2. A summary of Data Security on WINGS Systems

Data Security on WINGS Systems		
Devices	Network communication	Protocols

This project has received funding from the European Union's
Horizon Europe Framework Programme under grant agreement Nº 101057497.

AES (Advanced Encryption Standard) for:		NB-IoT:		UDP (User Datagram Protocol)
To safeguard the data stored on the device.	the confidentiality and integrity of the stored data even if an unauthorised individual gains access to the device.	impedes the ability of malicious entities to eavesdrop on or intercept the transmitted data	better communication capabilities with the devices	maintenance of the integrity of the communication
				uncorrupted data
				elimination of potential information loss
		Alternatives: 3G/4G/5G LoRA WiFi		

5.1.2.5 Data communication from sensor to gateway

AIRWINGS devices are compact by design and include sensors, communication module, power module, PCB, and any other hardware element that is needed (all inside the same external case/box). The communication between the sensors and the communication module takes place inside the AIRWINGS station with AES being implemented. The external case is protected and secure for operating in external environments (and certified for this). Additionally, there is a mechanism in place for alerting, in case someone attempts to open the external case.

Especially, for this case, a sensor is installed and monitors whether the box is opened or closed, avoiding in this way, anyone attempting to intervene at the device level and at the communication with the sensors.

This project has received funding from the European Union's
Horizon Europe Framework Programme under grant agreement № 101057497.

5.1.2.6 Data communication within the Cloud

AIRWINGS devices can communicate with the Cloud using several communication technologies (NB-IoT, LoRa, GPRS, 4G/5G, WiFi). Depending on the requirements (device's behaviour, power requirements, telecommunication costs) the most appropriate communication technology is selected. Similarly, the most appropriate communication protocol is implemented (e.g. UDP packets). It should be mentioned that in all cases security measures are applied. Either these are already being supported by the communication technology that is selected or they are programmed at the device level.

In the first case, for example, NB-IoT is considered as secure by design, impeding the ability of malicious entities to eavesdrop on or intercept the transmitted data. Similarly, security is provided from the other communication technologies as explained in Section 1.1.6.2. Moreover, AIRWINGS devices are also programmed to be secure and to guarantee secure exchange of messages, utilising encryption standards like WPA2 for secure data transmission. Details regarding security at protocol level (for communication with the Cloud) is provided in Section 0.

5.1.2.7 Data communication from the Cloud to the user interface or external systems

AIRWINGS system (at the platform level) supports data communication with the user interface, but also, data communication with external systems. This communication is mainly offered through the API of the cloud platform and in all cases, security is considered in advance. Therefore, for retrieving data from the API, partners are required to use HTTPS requests, which are secured with API token validation. This method ensures that only authenticated users can access the data, maintaining the integrity and security of the information exchanged. The API token is a unique identifier assigned to each partner, which must be included in the HTTPS headers of each request to authenticate and authorise the access.

5.1.3 Measurement and monitoring implemented by Thinnect

5.1.3.1 General Communication Architecture Design

The Thinnect EDIAQI IoT solution follows a typical IoT architecture that consists of autonomous wireless devices at the network edge delivering collecting and sending data to a central Cloud platform. The Thinnect EDIAQI solution employs also an Edge gateway that aggregates data from all the sensors in a building, forwarding it to the Cloud platform. The solution architecture is depicted below on Figure 3. For sensor communication a mesh networking technology is employed, which enables deployment of even a high number of sensors with low cost and low configuration overhead.

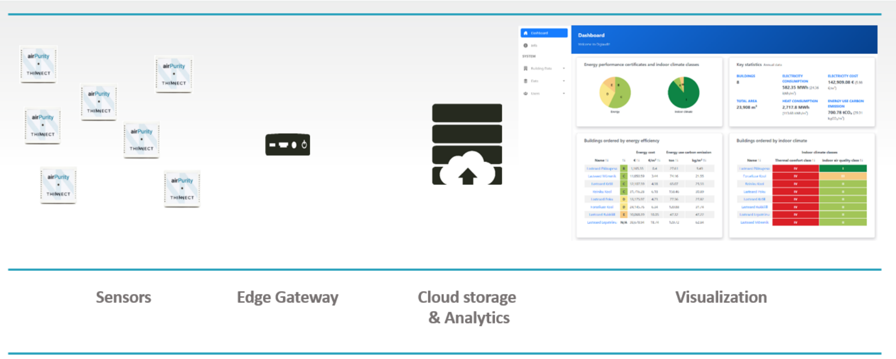

Figure 3 – Thinnect IoT sensing solution employing an Edge gateway

The benefits of a gateway-based solution that employs a mesh network for sensor communications are:

1) Higher flexibility for deployment and lack of need for creating a network infrastructure or an existing network infrastructure compared to WiFi-based solutions.
2) Lower investment cost and lower operational cost compared to cellular-based solutions.

The Thinnect EDIAQI solution also makes use of Mist computing architecture, where the intelligence is pushed to the network edge, enabling remote management of the Edge functionality. Mist computing complements Fog computing, where intelligence is located at the network edge in the gateways and Cloud computing, where intelligence is in the Cloud. With Mist computing the data can be pre-processed at the Edge nodes, aggregating, averaging and filtering data if needed. Also, intelligent data communication triggers can be put in place, which force immediate data communication when certain parameters are out

This project has received funding from the European Union's
Horizon Europe Framework Programme under grant agreement Nº 101057497.

of pre-defined bounds. Mist computing makes it possible to reduce the amount of data communicated from the sensors to the Cloud as only the data that needs to be communicated is communicated. This, in turn, reduces the power consumption of the devices if the software is implemented properly, enabling battery-powered sensing devices with a long battery lifetime.

5.1.3.2 Data Communication from Sensor to Gateway

Sensor solutions used in pilots that utilise technology from Thinnect can make use of Thinnect's unique Edge Public Key Infrastructure (PKI) technology, which ensures that every sensor can be uniquely authenticated and authorised. The gateway connection is secured using standard and proven Internet security methods, applying TLS and SSH.

5.1.3.3 Data Communication from Gateway to Cloud

The data communication from the gateways to the Cloud is managed by the gateway – in each gateway the connection is established to the Cloud backend. Methods like the TLS protocol ensure secure communication from source to recipient (end-to-end) over public networks, providing authentication, data integrity and confidentiality.

5.1.3.4 Data Communication inside the Cloud

Inside the Cloud the data is secured by standard Cloud security approaches – the upstream inbound connections (gateways to Cloud) are only enabled for known devices utilising secure connections with SSH or VPN standard authentication using certificates following the TLS standard.

5.1.3.5 Data Communication from Cloud to UI and external systems

The connection used by the data users to the Cloud data storage employs proven and standard Internet security methods. User authentication is implemented using tokens. A token authentication means that each request to a server is accompanied by a signed token which the server verifies for authenticity and only then responds to the request. Tokens are used because tokens are stateless, tokens can be generated from anywhere, and finally, tokens have fine-grained access control.

User authentication is implemented using HTTP realm login where log in at /token with username and password yields a token to be used as a username on further requests. The token is not auto renewed, its lifetime is set by default to 4 hours.

User access rights are controlled via user roles where each role has a set of detailed read-write privileges for specific functionality. Endpoints specify the access rights they need.

From the Thinnect EDIAQI server the outbound API connections are initiated by the Thinnect server and the connections are secured using the security methods applied at the receiving server.

5.2 FROST Server Security

Commented [JP1]: Input is needed for this section.
Alessandro, LAS will provide input on their server.
Heimo, KNOW will provide input on server to server and FROST API security.
Piergiorgio, Deda can provide input on general FROST security.

The EDIAQI project uses the Frost server to collect data from individual platforms to a single server. In this chapter the FROST server security measures are described.

5.2.1 General Frost Server Security Measures

The Frost server implements the security measures generally applicable for the FROST servers.

5.2.2 Frost Server Security Measures Implemented in EDIAQI

In EDIAQI, the standard role-based FROST server security measures are implemented to ensure confidentiality and privacy of data.

5.2.3 Frost Server to Frost Server Security Measures Implemented in EDIAQI

As EDIAQI implements multiple instances of Frost servers, server to server security measures had to be implemented. The server-to-server security measure apply also standard role-based FROST server security measures to ensure data confidentiality and privacy.

5.3 Overall assessment of applied security measures

The overall assessment of the security measures applied by EDIAQI technology providers is positive. All the indoor climate data collection solutions used in EDIAQI apply a reasonable level of data security for data collection and processing. In cases of some systems the end-

This project has received funding from the European Union's
Horizon Europe Framework Programme under grant agreement Nº 101057497.

to-end data security of the solution must be still evaluated to ensure that end-to-end data privacy and security can be maintained.

5.4 Recommendations on security measures

No major recommendations for improvement of security measures are applicable to any of the data collection solution technology providers. As discussed, the end-to-end data security and privacy aspects must be evaluated during the project to ensure that proper security measures are applied. The findings of these evaluations will be included in Deliverable 4.9. During the project recommendations will be provided to the project participants on an ongoing basis.

5.5 Data anonymisation

In EDIAQI the data collected on environmental parameters provides information on the indoor climate quality, which may be considered as sensitive information by the organizations that own, manage and maintain the buildings. This data may however also reveal information about the usage patterns of the buildings and the activities of the individuals. This information could be used maliciously by individuals and organisations for unethical business practices as well as for criminal activities. As the collected data should be made available to a research community for analysis, anonymisation should be used to protect the privacy.

Anonymisation cannot be applied to the collected sensor data as it would render the data useless from the research perspective – it can't be used as an input for research. Anonymisation cannot be well applied for the temporal aspects of data either as in this case the usability of the data for research purposes may be also hindered. Therefore, the data anonymisation can be applied to the locations of the data, eliminating the possibility that the data are associated with specific individuals or organisations.

The anonymisation should be applied to the data at the Cloud level and below the data elements used in EDIAQI that may need data anonymisation are listed with a recommendation on the need for anonymisation for each data element.

Table 3 Anonymisation of EDIAQI data in STA data model: “Locations” entity

STA element	Description (EDIAQI)	Data type	Example	Anonymization advisable
name	Name or code of the location expressed as: - name of the building (if available) - name or code of the single room or building part	String	Istituto Copernico-Carpeggiani Ferrara, classe 6A	YES
description	Description of the location where the thing is located, with verbose details	String	The thing has been installed at the ground floor in classroom 6A of the high school Istituto Copernico-Carpeggiani	YES
encodingType	Type of encoding for representing the location of the pilot building	Default = ‘application/geo+json’[4]	application/geo+json	NO
location	(note: the following rows represent an excerpt of a normal GeoJSON structure for a point representing the exact location as longitude/latitude of the thing installed)			YES
type	Type of spatial feature representing the location of the thing	Default = 'Point'	Point	NO
coordinates	List of coordinates	List of geographical coordinates		YES
0	First coordinate (long) of the location of the pilot building	Double (min. 5 decimals, max 8)	11.64273	YES

[4] Currently geojson is the only Location encodingType of the SensorThings API. GeoJSON supports the following geometry types: Point, LineString, Polygon, MultiPoint, MultiLineString, and MultiPolygon. Geometric objects with additional properties are Feature objects. Sets of features are contained by FeatureCollection objects: https://datatracker.ietf.org/doc/html/rfc7946

This project has received funding from the European Union’s
Horizon Europe Framework Programme under grant agreement Nº 101057497.

1	Second coordinate (lat) of the location of the pilot building	Double (min. 5 decimals, max 8)	44.82996	YES
properties	A JSON Object containing user-annotated properties as key-value pairs	JSON Object		YES
identifier	Project id for the location with the following pattern: - EDIAQI (fixed) - 2-digit code for pilot area (es. FE) - 3-digit code for the location of the thing	String	EDIAQI.FE.001	NO
building	General information about the pilot building where the thing is located			YES
owner	URI of the owner of the pilot building	URL	https://www.provincia.fe.it/	YES
identifier	Local identifier of the pilot building	String	FEIS01200X	YES
address	Address locator, expressed as human readable designator or name that allows a user or application to reference and distinguish the address from neighbour addresses, within the scope of a thoroughfare name, address area name, administrative unit name or postal designator, in which the address is situated	String	Via Pontegradella 25, 44123 Ferrara	YES
country	Code of the country (ISO 3166-1) where	String	IT	NO

	the pilot building is located			
use	Current use of the whole building, according to INSPIRE Directive (2007/2/CE) codelist for “Buildings” data theme[5]	Codelist Residential - Collective residence - Individual residence - Residence for communities Industrial Commerce and services - Office - Public service - Trade	Public service	NO
use_details	Detailed information about the use of the whole pilot building	String	Public high school	YES
3Dmodel	External reference to 3D model representing the whole building as geographical dataset (e.g. in CityGML or KML format)	URL		YES
room	Details about the placement of the thing (sensor placement in §**Error! Reference source not found.**)			NO
identifier	Local identifier of the room	String	6A	YES
elevation	Overall height above sea level (meters) of the thing	Double (2 decimals)	14.23	NO
level	Floor number where the thing is located (single room or building part)	Integer	3	NO
height	Height from room floor (meters) of the thing	Double (2 decimals)	1.55	NO

[5] https://inspire.ec.europa.eu/codelist/CurrentUseValue

d_ceiling	Distance from room ceiling (meters) of the thing	Double (2 decimals)	3.12	NO
d_horizontal	Horizontal distance to adjacent walls (meters) of the thing	Double (2 decimals)	0.90	NO
3Dmodel	External reference to 3D model representing the single room (e.g. in CityGML or KML format)	URL		YES

This project has received funding from the European Union's
Horizon Europe Framework Programme under grant agreement № 101057497.

Table 4 Anonymisation of EDIAQI data in STA data model: “FeatureOfInterest” entity

STA element	Description (EDIAQI)	Data type	Example	Anonymization advisable
name	Name or code of the single room or building part that is monitored	String	Class6A	YES
description	Verbose description of the feature of interest (single room or building part)	String	Classroom 6A	YES
encodingType	Type of encoding for representing the feature of interest (single room or building part)	Default = ‘application/geo+json’[6]	application/geo+json	YES
feature	(note: the following rows represent an excerpt of a normal GeoJSON structure for a polygon with N vertexes)			NO
type	Geographical extent of the area (single room or building part) whose properties are being measured	Default = 'Polygon'	Polygon	NO
coordinates	List of coordinates	List of geographical coordinates		NO
0	1st vertex of the polygon			NO
0	Longitude of the feature of interest (single room or building part)	Double (min. 5 decimals, max 8)	11.64273	NO
1	Latitude of the feature of interest (single room or building part)	Double (min. 5 decimals, max 8)	44.82996	NO
1	2nd vertex of the polygon			NO
0	Longitude of the feature of interest (single room or building part)	Double (min. 5 decimals, max 8)	11.64273	NO
1	Latitude of the feature of interest (single room or building part)	Double (min. 5 decimals, max 8)	44.82996	NO
N	Nth vertex of the polygon			NO
0	Longitude of the feature of interest (single room or building part)	Double (min. 5 decimals, max 8)	11.64273	YES
1	Latitude of the feature of interest (single room or building part)	Double (min. 5 decimals, max 8)	44.82996	YES
properties	A JSON Object containing user-annotated properties as key-value pairs	JSON Object		NO

identifier	Project id for the feature of interest, with the following pattern: - EDIAQI (fixed) - 2-digit code for pilot area (es. FE) - 3-digit code for the feature of interest	String	EDIAQI.FE.001	NO
room	Details about the placement of the thing (sensor placement in §**Error! Reference source not found.**)			NO
identifier	Local identifier of the single room or building part	String	6A	NO
elevation	Overall height above sea level (meters) of the thing	Double (2 decimals)	14.23	NO
level	Floor number where the thing is located (single room or building part)	Integer	0	NO
3Dmodel	External reference to 3D model representing the single room (e.g. in CityGML or KML format)	URL		NO
use	Current use of the feature of interest (single room or building part)	Codelist - Living room - Bedroom - Kitchen - Open space (kitchen and living room) - Office - Laboratory - Classroom - Sport facility - Restaurant - Subway station	Classroom	NO
use_details	Detailed information about the use of the feature of interest (single room or building part)	String	Classroom used by teachers and students for daily lessons	NO
floor_area	Area (in m^2) of the floor of the feature of interest (single room or building part)	Integer	120	NO
glazed_area	Area (in m^2) of the glazed surface (windows) of the feature of interest (single room or building part)	Integer	15	NO
ventilation	Presence of ventilation system in the single room or building part	Codelist: - None - Manual - Electric fan - Air conditioning system - Air purifier	Air conditioning system	NO

This project has received funding from the European Union's Horizon Europe Framework Programme under grant agreement № 101057497.

		- Central ventilation meeting		
occupants[7]	Information about the occupants of the single room or building part			NO
number	Number of occupants	Integer	22	NO
type	Typology of occupants	Codelist: - Kindergartners - Students - Teachers - Patients - Visitors - Workers - Household dwellers	Students	NO
gender	Gender of occupants (per typology)	Codelist: - Female - Male - All genders - Irrelevant	All genders	NO
age	Age of occupants (per typology)	Codelist: - under 5 years old - 5-10 - 10-18 - 18-25 - 25-65 - over 65 years old	10-18	NO

[7] The element "occupants" can be repeated more than once, with different typologies.

This project has received funding from the European Union's
Horizon Europe Framework Programme under grant agreement № 101057497.

6. Status and next steps

While the state of data security in EDIAQI is satisfactory (after some improvements have been applied), it cannot be guaranteed that data privacy is ensured at a satisfactory level in the long term with the current approaches applied in EDIAQI. Therefore, the recommendations from this report should be applied to the data collected in EDIAQI to ensure the desired level of data privacy and security. The project partners are encouraged to apply the recommendations from this report in their solutions and data delivery methods to ensure data privacy.

6.1 Next steps for Deliverable D4.9

The next iteration of the EDIAQI security report – Deliverable D4.9 will evaluate the changes implemented and measures applied during the project to ensure proper long-term data privacy and security. Throughout the course of the project in Task 4.6 the data privacy and security aspects will be monitored, and the project partners will be consulted in these topics. The next report will also address important security topics like

- Authentication and authorization methods like Multi-Factor Authentication (MFA), including role-based access control.
- Security for API interfaces.
- Incident and intrusion detection methods, potentially tools for monitoring of security aspects such as system activities, network traffic and user interactions.
- Continuous monitoring and incident response for EDIAQI systems.
- Security tools and methods applied, such as firewalls.
- Management of security updates and fixes to system components at different levels.
- Data backup and recovery methods where applicable.

The involved project partners are considering holding a small-scale cyber security exercise to assess the security of the EDIAQI solution components and to teach relevant parties about cyber security methods. In the context of the possible cyber security exercise EDIAQI solution components would be set up and attacked by red team members with the goal to identify vulnerabilities in the system(s) and their components.

Deliverable D4.6

Privacy and IoT Security – Version 1

Work Package 4
MONITOR
Version: Final

This project has received funding from the European Union's
Horizon Europe Framework Programme under grant agreement Nº 101057497.