COMPAIR

DELIVERABLE

D4.7 Citizen Science IoT Standardisation White Paper

Project Acronym:	**COMPAIR**	
Project title:	Community Observation Measurement & Participation in AIR Science	
Grant Agreement No.	101036563	
Website:	www.wecompair.eu	
Version:	1.0	
Date:	28 June 2024	
Responsible Partner:	DV	
Contributing Partners:	ATC, ECSA, IMEC, TELR, SODAQ,	
Reviewers:	Aouefa Amoussouvi (ECSA) Paul Sorrell (ECSA) Andrew Stott Otakar Cerba (UWB)	
Dissemination Level:	Public	X
	Confidential, only for members of the consortium (including the Commission Services)	

This project has received financial support from the European Union's Horizon 2020 Programme under grant agreement no. 101036563

Revision History

Version	Date	Author	Organisation	Description
0.1	12.03.24	Lieven Raes	DV	Initial structure
0.2	07.06.24	Lieven Raes	DV	Content chapters 4-8
0.3	13.06.24	Aouefa Amoussouvi, Paul Sorrel	ECSA	Content chapter 3.1.
0.4	14.06.24	Pavel Kogut	21C	Examples of innovative IoT initiatives for section 3.1
0.9	19.06.24	Otakar Čerba, Andrew Stott	DV	Review
0.9	27.06.24	Aouefa Amoussouvi	ECSA	Review
1.0	28.06.24	Lieven Raes	DV	Final version

Table of Contents

List of Figures

1. Executive Summary

In contemporary urban environments, a diverse array of Internet of Things (IoT) devices are employed to monitor physical conditions and urban activities. However, only a fraction of the available IoT resources are utilised for decision-making processes. Achieving interoperability at technical, semantic, and organisational levels is essential to fully leverage the potential of diverse IoT device networks and types.

This white paper is about how IoT interoperability can be achieved to facilitate decision-making in a Citizen Science context and how results can be used to enrich e.g. new policy instruments like digital twins with relevant live (local) data. Aside from the IoT standards on a network and semantics level, overall system architecture in a smart city and digital twin context is important, too. Such architectures define how data can be described, managed and stored to generate, e.g. qualitative time-series data, feeding into data dashboards, digital twins and simulation models to enable, e.g. predictive modelling.

Also, the implementation of standard communication protocols and robust infrastructures is critical for the seamless integration of IoT in increasingly data-driven governance frameworks. At a European level, the creation of a level playing field for IoT devices with low power and bandwidth connectivity is key for the success of large IoT Citizen Science projects. Today, a reliable network is missing across Europe.

This white paper offers a comprehensive overview of relevant IoT standards applicable across various applications. It details the interoperability requirements for different sensor types, encompassing both professional high-fidelity IoT networks and community-driven citizen science initiatives by harnessing new data sources and tools, such as visualisation dashboards and digital twins. Decision-makers can gain better insights and conduct more holistic policy impact assessments. This white paper includes several case studies that illustrate the potential of IoT to develop data-driven solutions for urban environments and their inhabitants.

2. Introduction

The use of IoT devices in Citizen Science projects is growing. IoT devices add new possibilities like near-live transmission and processing of results, on-device data processing, availability checking, and automated software updating and maintenance, to name a few. IoT-based Citizen Science also comes with a number of challenges, like data processing, quality control, and data sharing.

COMPAIR is unique in the way it uses multiple sensor devices. This means COMPAIR was also confronted with issues like data exchange and integration. Standardisation is even more important here to integrate and expose data using open, well-known IoT software and hardware (transmission) standards.

This white paper starts with a brief history of the use of IoT in Citizen science, including several examples from the COMPAIR pilot member states, such as Belgium, Bulgaria, and Germany (Chapter 3).

Chapter 4 pinpoints how IoT devices fit into a layered ICT architecture as used in, e.g. Digital Twins and other Smart-City-related platform solutions.

Chapter 5 focuses on the types and characteristics of IoT devices in the broad smart-city domain. These characteristics are related to device capabilities, connectivity, data processing, and metadata.

Chapter 6 is about IoT data standards and semantics. It focuses on the IoT standards themselves, IoT related to Digital Twins, and their relation to the OASC Minimal Interoperability Mechanisms.

Chapter 7 discusses how IoT sensors, including networks, that are part of a Citizen Science project can be part of a Digital Twin solution.

Chapter 8 discusses an IoT and standardisation issue related to the lack of a level playing field in Europe regarding LTE-M, LTE-MTC, and NB-IoT low-power and low-bandwidth network availability. This chapter is based on a note shared and discussed with the European Commission during the COMPAIR project.

3. The history of the use of IoT data in Citizen Science projects

3.1. History of Citizen Science Initiatives

The EU Horizon 2020 and Horizon Europe funding programmes for research and innovation have significantly advanced transnational citizen science initiatives across Europe. Horizon 2020 (2014-2020), with a budget of nearly €80 billion, and Horizon Europe (until 2027), with a budget of €95.5 billion, have promoted cross-border collaboration and fostered innovative projects addressing regional and global challenges[1][2]. By providing substantial financial support, these programmes have enabled diverse communities to engage in scientific research, enhancing data quality and broadening the impact of European citizen science efforts in the last decades. Additionally, technological advancements from these initiatives, driven by the diverse project locations, have led to the creation and deployment of more sophisticated data collection tools and platforms, significantly improving the efficiency and reach of citizen science.

Early Citizen Science History

Citizen science, the practice of public participation and collaboration in scientific research, has a storied history that dates back centuries before the conception of a professional scientist entered into common practice and understanding (Miller-Rushing et al., 2012). Initially informal and often conducted by amateur naturalists and hobbyists, citizen science has transformed significantly over time to encompass research in fields spanning the scientific spectrum, including air and water quality, biodiversity, public health, climate change, sustainability, artificial intelligence and more. Recent advancements in technology, particularly the advent of the Internet of Things (IoT) and connected objects, have revolutionised citizen science, expanding its scope of possibilities and making it more accessible, efficient and impactful (Vohland et al., 2021).

The 20th century saw the formalisation of citizen science, particularly in fields like botany, animal biodiversity, weather phenomenology and astronomy. Ornithology programs such as the Audubon Society's Christmas Bird Count, initiated in 1900, exemplify early structured citizen science projects[3]. These programs relied on the public to collect data on bird populations, which was then analysed by scientists. This era marked the beginning of systematic data collection and the realisation of the immense potential of leveraging public participation in scientific research.

Though there is no single agreed upon definition of citizen science, an early use of the term can be traced back to an article 'Lab for the Environment' in the January 1989 issue of MIT Technology Review featuring three community-based laboratories studying environmental issues (Kerson, 1989). The article describes how the nature conservation organisation 'Audubon Society' recruited 225 volunteers in a 5-week "citizen science" program" to collect acid-rain data from all 50 states of the US. A few years later, Alan Irwin in the United Kingdom and Rick Bonney in the United States independently conceptualised and published definitions of citizen science in 1995 and 1996, respectively. Bonney's work at the Cornell Lab of Ornithology emphasised public engagement in scientific research, particularly in ornithology, while Irwin's research focused on the relationship

[1] https://wayback.archive-it.org/12090/20220124080607/https://ec.europa.eu/programmes/horizon2020/what-horizon-2020#Article

[2] https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe_en

[3] https://www.audubon.org/community-science/christmas-bird-count

between science and society, advocating for greater public involvement in scientific endeavours (Bonney, 1996).

Citizen Science Technological Advancements with the Digital Age

The advent of the internet marked a pivotal moment in the evolution of citizen science. Online platforms enabled broader participation, allowing people from diverse geographic and demographic backgrounds to contribute to scientific research. Projects like SETI@home, launched in 1999, harnessed the power of distributed computing, inviting volunteers to use their home computers to analyse radio signals for signs of extraterrestrial intelligence[4]. As digital technology continued to evolve, the scope and scale of citizen science projects expanded. Online databases, mobile applications, and social media platforms emerged as powerful tools for data collection, dissemination, and community engagement. The creation of platforms like Association for Advancing Participatory Sciences, Zooniverse, SciStarter, European Citizen Science Platform, which hosts a multitude of citizen science projects, further democratised access to scientific research[5][6][7][8].

IoT and Connected Objects in Citizen Science

The integration of IoT and connected objects into citizen science has significantly enhanced its capabilities to explore questions in areas previously inaccessible due to technological limitations. IoT refers to the network of physical devices embedded with sensors, software, and other technologies to connect and exchange data with other devices and systems over the internet (Newman et al., 2012).

IoT devices, such as environmental sensors, wearable technology, and smart home devices, have transformed how data is collected and analysed in citizen science projects. These devices can continuously monitor and transmit data in real-time, providing more accurate and comprehensive datasets than ever before. For instance, environmental monitoring projects, like COMPAIR, now utilise IoT-enabled sensors to track air quality and share their data through an online platform. Such projects provide valuable data for understanding and addressing pollution and climate change as well as foster a sense of community and collective action[9].

3.2. Innovative IoT-based Citizen Science Projects and Initiatives, an overview

There are numerous IoT-based Citizen Science Projects and Initiatives running using air quality sensors. The InfluencAir[10] project supported by the Open Knowledge community, the MeterAC[11] project in Bulgaria, an area that is less developed on Citizen Science, and the world wide senseBox[12] project are three examples of innovative IoT based Citizen Science initiatives. It must be noted that almost all IoT Citizen Science projects are using just one type of (static) sensor. Some initiatives like senseBox are using multiple sensors.

[4] https://setiathome.berkeley.edu/
[5] https://participatorysciences.org/
[6] https://www.zooniverse.org/
[7] https://scistarter.org/
[8] https://eu-citizen.science/
[9] https://www.wecompair.eu/who-we-are
[10] https://www.facebook.com/Influencair/
[11] https://my.meter.ac/index.html
[12] https://sensebox.de/

The InfluencAir project set up a network of air monitoring devices hosted by citizens to measure air pollution in and around Brussels. Initially, measurements were made using the SDS011 sensor combined with the WiFi microchip NodeMCU ESP8266. Some tests were also performed with the Honeywell HPMA115S0 sensor. However, the project needed a technology capable of covering several kilometres of outdoor space in the Brussels region. The Low Power Wide Area Network (LoRaWAN) was chosen for data transmission instead of WiFi which has a shorter range. The team developed their sensor kit based on the LoRaWAN technology, which helped to cover a larger area by allowing low-powered devices to communicate with other nodes in a network over long-range wireless connections (Open Knowledge Belgium, 2018).

METER.AC dataset, a comprehensive and accessible resource, monitors atmospheric pressure, temperature, relative humidity, particulate matter, and background radiation in over 100 locations throughout Bulgaria. The measurements are conducted by low-power, maintenance-free nodes equipped with common hardware and software components. The data flow commences from the sensor layer operating in the physical environment, followed by a media layer for internet communication (Terziyski et al., 2020).

senseBox is an easy-to-use IoT toolkit which includes different sensors for measuring air pollution (PM10) and atmospheric conditions (temperature, humidity, air pressure, illuminance, UV radiation). Developed in Germany, around 10,000 senseBoxes are currently installed worldwide. Data from the senseBox to the openSenseMap can be transferred via WiFi, Ethernet or LoRa. The WiFi connection is based on the ATWINC1500 microchip by Atmel, which has a very low energy consumption and a long range. With the Ethernet option, data from senseBox is transmitted using the ethernet cable to the router; the method is based on the W5500 Mikrochip by Wiznet which allows a high ethernet data transmission rate. Finally, the LoRa interface provides a low-energy, cost-free option to upload data using the LoRa-Radio-Standard. Existing LoRa-Networks such as TheThingsNetwork are used for data transmission. The community of TheThingsNetwork provides the necessary infrastructure and is available in more and more regions.

Today, three types of data transfer solution families are mainly used: LoRa, WiFi, and wired Ethernet connectivity. LoRa ("Long Range") is employed in the context of a low-power wide-area network based on wireless telecommunication designed to enable long-range communications at a low bit-rate among low-power connected objects, such as sensors running on a battery (ibid.).

4. Role of ICT Architectures and Standardisation

The application of the Internet of Things (IoT) plays an important role in various smart city domains, including smart agriculture, city services, energy, health, home, industry, infrastructure, and transport. IoT serves as a critical enabler by continuously translating information from the physical world into the digital realm. This digitised information can then be processed into knowledge pertinent to one or more of the domains above, thereby supporting decision-making and policy formulation. Local Digital Twins (LDT) are considered one of the key technologies that will support the use of IoT in the decision-making process. A local Digital Twin can be described as follows:

> A **Local Digital Twin** is a **virtual representation** of the **physical assets**, **processes**, and **urban intelligence** within a **geographically located community**, which reflect and derive from **cross-sectorial**, **historical, and (near) real-time data**. Its purpose is to enhance **evidence-informed decision-making** according to **ethical standards and principles** at the **operational, strategic, and tactical levels** to better meet communities' **needs**. LDTs support predictive simulation modelling, combining multiple technologies offering **open and interoperable components**, allowing **integration with legacy systems** and other components in a **federated architecture**.

Regarding Information and Communication Technology (ICT) architecture and standardisation, IoT shares several characteristics with the multi-layered architecture of LDTs. Syed et al. (2021) outline a five-layer architectural model, which includes a sensing layer comprised of domain-specific sensors, actuators, and mobile elements. Data from this layer is transmitted to a middleware layer via a network layer that spans various network technologies and topologies (Figure 1). The middleware layer offers a generic open application interface (API) and manages data to provide a comprehensive array of services suitable for use in LDTs.

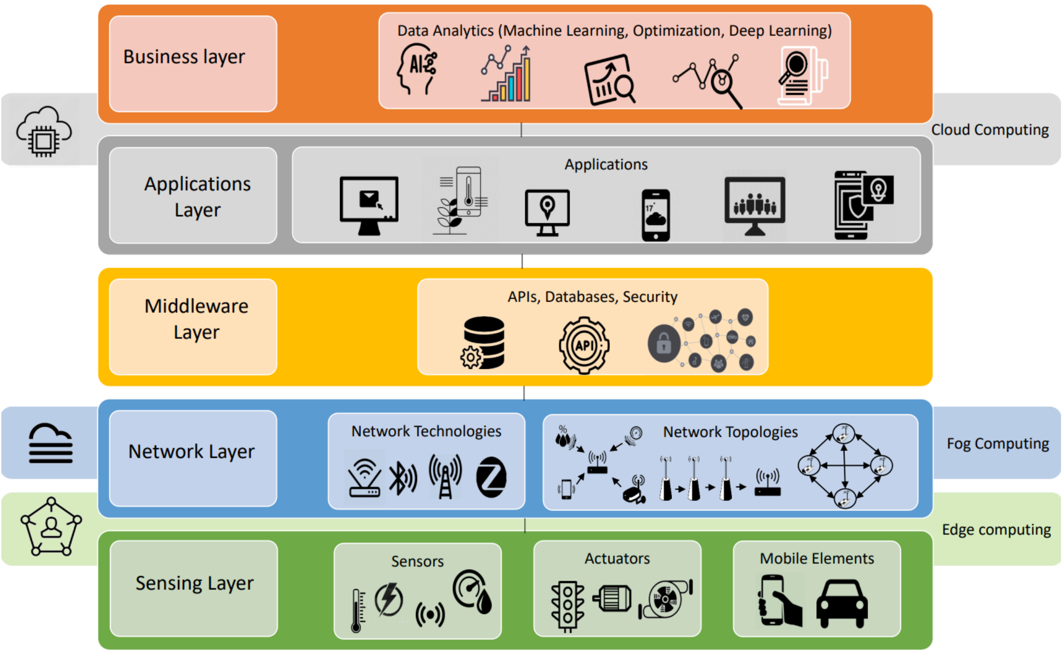

Figure 1:The place of IoT in a layered ICT architecture as e.g. used in Digital Twins (Syed et al., 2021)

The increasing affordability of IoT sensor technology and the accessibility of network technology have facilitated the development of denser IoT sensor networks, which can deliver more localised city data and subsequently create new opportunities for policy support. Furthermore, decision-making driven by IoT and Local Digital Twins (LDTs) can leverage the novel possibilities presented by citizen science. In Europe, citizen science initiatives are advancing the state of the art. For instance, the COMPAIR project integrates science, technology, and co-innovation to enhance the role of community-based IoT networks in local policymaking. This initiative involves the development of various sensors and applications designed to empower urban stakeholders to address climate change and air pollution.

Beyond the scope of citizen science, IoT contributes to geospatial and time-bound measurements that can be archived as geo-time series (Carfantan et al., 2020). The spatiotemporal aspect of IoT provides a crucial context for integrating data from diverse sensors and data streams into an LDT. Spatial and temporal reasoning, based on an ontology of space as described by Stock in the late 1990s (Stock, 1997), has become integral to almost all effective IoT standards within the smart city context. While real-time sensor data may be less critical for long-term policy decisions, it is essential for operational management and serves as a significant input for Integrated Command and Control Centers (ICCCs).

Chapter 5 delves into the types and characteristics of IoT sources, data standards and semantics, the role of IoT sensor networks within an LDT, and several practical use cases of IoT sensor-based applications related to LDTs and the CitiVerse.

5. Types and characteristics of IoT devices

5.1. IoT Device functionalities and characteristics

In the context of smart cities and LDTs, IoT devices play a crucial role in collecting and exchanging data to enable intelligent and efficient urban management. These devices can be categorised into various types based on their functions and applications (Rp et al., 2021), as described below.

Sensors for monitoring physical conditions and chemical substances

- *Environmental sensors:* measure parameters such as air quality, temperature, humidity, and pollution levels;
- *Noise sensors:* monitor noise pollution levels in different areas of the city;
- *Light sensors:* control street lighting based on ambient light conditions;
- *Water quality sensors:* monitor the quality of water sources and detect contaminants.

Infrastructure monitoring devices

- *Structural health monitoring sensors:* monitor the condition of bridges, buildings, and other infrastructure;
- *Smart grid sensors:* monitor and manage the electrical grid to optimise energy distribution;
- *Traffic monitoring sensors:* collect data on traffic flow, congestion, and vehicle movements.

Public safety and security devices

- *Surveillance cameras:* monitor public spaces for safety and security;
- *Explosion detection systems:* detect and locate explosions to improve emergency response;
- *Smart streetlights with cameras:* combine lighting with video surveillance for enhanced security.

Transportation and mobility devices

- *Smart parking sensors:* help drivers find available parking spaces and optimise parking management;
- *Connected vehicles:* vehicles equipped with IoT technology for real-time communication with each other and infrastructure;
- *Public transportation trackers:* provide real-time information on the location and status of public transport.

Waste management devices

- *Smart bins:* monitor waste levels and optimise waste collection routes;
- *Environmental sensors for landfills:* monitor the environmental impact of landfills.

Health monitoring devices

- *Health wearables:* collect health data from citizens for public health analysis;
- *Ambient assisted living devices*: assist in monitoring the well-being of the elderly or individuals with special needs.

Communication devices

- *IoT Gateways:* Facilitate communication between various IoT devices and the central network;
- *Communication Nodes:* Enable device-to-device communication in a mesh network.

Information and display devices

- *Smart kiosks:* provide information on local services, events, and wayfinding;
- *Digital signage:* display real-time information and advertisements in public spaces.

IoT devices are deployed across a wide range of environments, each tailored to meet specific needs. While common characteristics unify them, the unique demands of smart homes, factories, containers, and smart cities shape their specific attributes.

All IoT devices share the fundamental trait of connectivity, enabling seamless communication and creating an interconnected ecosystem. At their core, these devices collect and process data through embedded sensors. Whether it's environmental conditions, machine status, or user behaviour, IoT devices are designed to capture context-relevant information.

Data transmission is the lifeblood of IoT systems. Using protocols such as MQTT or CoAP, these devices communicate with central systems or other devices, ensuring a steady flow of information. Remote monitoring and control capabilities are ubiquitous, allowing centralised management by users or automated systems.

Security is a paramount concern across all IoT domains. Robust measures are implemented to secure data transmission, authenticate devices, and protect against unauthorised access, ensuring the trustworthiness of these interconnected systems.

In specific contexts, IoT devices exhibit distinctive characteristics:

- **Smart Homes**: The focus is on user comfort, convenience, and energy efficiency. Devices like smart thermostats, lighting systems, and security cameras integrate with voice assistants and mobile applications, offering users unprecedented control over their environments.

- **Smart Factories**: The emphasis is on industrial automation, process optimization, and efficiency. Industrial sensors, programmable logic controllers, and robotic systems work together to achieve high precision. These devices integrate with manufacturing execution systems, forming the backbone of modern manufacturing.

- **Containers and Logistics**: These devices track and monitor the location, condition, and security of goods during transportation. GPS trackers, temperature sensors, and security systems ensure the integrity of shipments. Integration with logistics management systems enables real-time tracking and route optimization.

- **Smart Cities**: Focusing on urban planning, sustainability, and public services, a diverse array of devices is deployed. Environmental sensors, traffic monitoring systems, and public safety devices contribute to a networked urban infrastructure. Integration with city management platforms and data analytics supports informed decision-making, paving the way for intelligent urban development.

In summary, while IoT devices share common traits like connectivity and data processing, their specific applications in various environments define their unique attributes, enhancing efficiency, security, and functionality across diverse domains.

5.2. Strengths and weaknesses of wired and wireless IoT devices

The mode of internet connectivity, whether wired or wireless, introduces significant variations beyond merely the method of linking to the internet. Devices connected to wired networks, such as Ethernet, exhibit inherent limitations in mobility, being confined to stationary settings due to their reliance on physical network infrastructure. The installation process for these devices involves routing physical cables, which can be time-consuming and require modifications to existing infrastructure. However, wired connections often provide notable stability and high-bandwidth connectivity once installed.

In contrast, wireless IoT devices, utilising technologies such as Wi-Fi, SIM cards, or narrowband, offer increased flexibility and mobility. Unconstrained by physical cables, these devices can operate in diverse locations. Installation is generally more agile and less encumbered by the complexities of physical wiring, although wireless performance can be affected by signal strength, interference, and network availability.

The difference in data transfer speed and bandwidth further distinguishes these connectivity modes. Wired connections typically deliver higher data transfer speeds and more consistent bandwidth, which is advantageous for applications requiring substantial throughput, such as industrial automation or video streaming. Conversely, wireless connections often contend with lower speeds, with bandwidth susceptible to variables such as signal strength and network congestion.

Power consumption is another critical consideration. Ethernet-connected devices draw power directly from the network infrastructure, providing an advantage for devices with strict power constraints. On the other hand, wireless devices, particularly those that rely on battery power, must balance energy consumption carefully. Wireless data transmission tends to be more power-intensive, impacting the overall battery life of these devices.

Security considerations also play a significant role. Wired connections are generally considered more secure due to their resistance to certain types of wireless attacks, with physical access to the network infrastructure providing an additional layer of protection. Wireless connections, however, are more vulnerable to eavesdropping and unauthorised access, necessitating robust security protocols such as WPA3 for Wi-Fi to protect IoT devices from potential threats.

Cost considerations add another dimension to this dichotomy. Wired infrastructure often incurs higher installation and maintenance costs due to the complexities of physical cabling and potential structural alterations. In contrast, wireless solutions tend to have lower installation costs and greater scalability, especially in scenarios where deploying physical cables is impractical or economically prohibitive.

The choice between wired and wireless IoT connectivity is a nuanced decision driven by the application's specific requirements. Many IoT implementations strategically integrate wired and wireless connections to balance performance and flexibility, highlighting the adaptability of the interconnected devices landscape.

Furthermore, the advent of new types of mobile sensors has revolutionised the IoT ecosystem. Mobile sensors, characterised by their portability and adaptability, introduce new methodologies for data collection. This innovation enables incorporating new IoT data sources, such as those emerging from citizen science initiatives, representing a pivotal shift in the IoT paradigm.

5.3. Principles and standards of IoT connectivity

Within the framework of a smart city digital twin, the significance of device-to-device communication is paramount. This communication enables real-time collaboration among a diverse array of IoT devices, allowing them to harmoniously exchange information and contribute to the dynamic responsiveness of the city simulation. The real-time exchange of information is not merely a technical detail; it is the core of a system designed to respond to the dynamic conditions of a city in real time. IoT devices communicate with various types of server infrastructure, including data spaces, using a range of communication protocols and technologies.

This interconnectedness ensures that data from multiple sources—such as environmental sensors, traffic monitors, and public safety devices—can be integrated and processed efficiently. The ability to share data in real-time allows the digital twin to accurately reflect the city's current state, enabling informed decision-making and timely interventions. Communication protocols such as MQTT, CoAP, and others play a crucial role in facilitating this seamless data exchange, while technologies like 5G, Wi-Fi, and LPWAN (Low Power Wide Area Network) provide the necessary connectivity.

By leveraging robust device-to-device communication, smart city digital twins can offer comprehensive, real-time insights into urban environments. This capability supports a wide range of applications, from optimising traffic flow and managing energy consumption to enhancing public safety and improving residents' overall quality of life. Thus, device-to-device communication is a foundational element in smart city digital twins' effective operation and continuous evolution.

Server communication

- REST (Representational State Transfer): A widely used architectural style for designing networked applications. IoT devices can interact with servers through RESTful APIs, exchanging data in a stateless manner using standard HTTP (Hypertext Transfer Protocol) methods.
- SOAP (Simple Object Access Protocol): A protocol for exchanging structured information in web services. While less common in modern IoT applications, some legacy systems may use SOAP for communication.
- Message Queues: Devices can communicate with servers using message queue systems like RabbitMQ or Apache Kafka, where messages are sent to a queue and consumed by the server asynchronously. Message queues are essential components for IoT networks to accommodate a large number of connected devices effectively.

Device-to-device communication

- Bluetooth: Common for short-range communication between IoT devices. Bluetooth Low Energy (BLE) is often used for energy-efficient communication.
- Zigbee and Z-Wave: Wireless protocols designed for low-power, short-range communication between devices in home automation and industrial settings.
- NFC (Near Field Communication): Enables close-range communication between devices, often used for contactless payments and data exchange.
- LoRaWAN (Long Range Wide Area Network): Suited for low-power, long-range communication in scenarios like smart agriculture or asset tracking.
- 6LoWPAN (IPv6 over Low-Power Wireless Personal Area Networks): Optimizes IPv6 for low-power, constrained devices in IoT applications.
- Mesh Networks: Devices can form mesh networks, allowing them to communicate with each other and relay messages within the network. This is often used in scenarios where direct communication with a central server may not be feasible.

Cloud communication

- MQTT (Message Queuing Telemetry Transport): A lightweight and efficient publish-subscribe messaging protocol commonly used for IoT. It enables devices to publish messages to a central server (broker) and subscribe to receive messages from other devices or the server.
- CoAP (Constrained Application Protocol): Designed for resource-constrained devices, CoAP is a lightweight protocol that allows devices to communicate with the cloud using HTTP-like methods such as GET, POST, PUT, and DELETE.
- HTTP/HTTPS (Hypertext Transfer Protocol/Secure): IoT devices can communicate with cloud servers using standard HTTP or its secure version, HTTPS. This is suitable for scenarios where web-based communication is required.
- WebSockets: Provides full-duplex communication channels over a single, long-lived connection. It is often used for real-time communication between IoT devices and cloud services.

In a typical IoT ecosystem, devices often utilise a combination of communication methods tailored to factors such as range, power consumption, data size, and application requirements. The selection of a communication protocol is dictated by the specific needs of the IoT deployment, taking into account considerations like latency, reliability, and energy efficiency. Ongoing standardisation efforts and industry best practices aim to ensure interoperability and seamless communication within this complex and diverse domain.

Efficiency and device-to-device communication are of particular importance (Pawar & Trivedi, 2018). Instead of overwhelming a central server, devices share relevant information among themselves, optimising network bandwidth and reducing latency. This collaborative decision-making process is not limited to centralised systems but is distributed across interconnected devices, promoting a nimble and responsive approach. Even if a central server encounters issues, the city's digital twin maintains its functionality as devices continue to communicate, ensuring the uninterrupted operation of essential functions. This decentralised approach enhances system reliability and resilience against potential disruptions.

Through communication, devices share insights into their status, resource consumption, and operational conditions. This interaction fosters a symbiotic relationship where resources such as energy, bandwidth, and processing power are dynamically allocated, contributing to the overall efficiency of the smart city. Furthermore, devices authenticate and verify each other's identity, establishing a digital barrier against unauthorised access and potential cyber threats. This layer of security is particularly crucial in scenarios involving critical infrastructure.

5.4. IoT data processing

IoT data processing can occur at various locations, each playing a crucial role in how e.g Digital Twins receive and manage the data. Generally, data can be handled and processed through cloud computing, fog computing, and edge computing. Each of these concepts refers to the point at which the initial data processing is performed. According to Syed et al. (2021), the edge computing model allows data processing via an ‘edge node’ that, in the case of IoT, can be done on the IoT ‘thing’ level itself. In the case of fog computing, the data processing takes place on the level of the network devices, such as a router). In cloud computing, data processing occurs within a centralised computer network where all data is aggregated. Each data processing model, including cloud computing, has distinct capabilities and limitations. Some IoT devices may utilise edge computing, while others perform all data processing in the cloud. Generally, cloud computing facilitates heterogeneous data management from various devices, albeit at a high network cost and with elevated privacy risks. However, it supports more complex decision-making processes. In contrast, edge computing processes data locally on the device, resulting in lower network costs and enhanced privacy by keeping data on the device. However, it is less capable of handling multi-domain, multi-sensor data processing.

Quality of IoT devices and data

The quality of data derived from an IoT device significantly influences its utility and role within decision-making processes. Various metrics can gauge data quality depending on the intended application. Within a digital twin context, three key elements are critical to evaluating IoT device quality:

1) Accuracy and Precision: Accuracy pertains to the proximity of a measured value to its true value, while precision denotes the consistency of multiple measurements of the same entity. These qualities are primarily influenced by the sensor's measurement principle. However, overall accuracy and precision can be enhanced through subsequent data processing steps:

- Data Validation: Before utilising data, it undergoes validation to assess its quality and integrity. Identification of defects or tampering in IoT sensors can prevent the propagation of inaccurate data.
- Data Cleaning: This process involves removing outlier or corrupted data points. IoT devices, employing simple sensing principles, may produce raw sensor output susceptible to stochastic errors, necessitating data cleaning.

- Data Calibration: Calibration involves comparing IoT device output to a known-quality standard, facilitating the development of parameterized models to refine raw measurements. Since IoT devices often rely on basic measurement principles sensitive to extraneous parameters, calibration necessitates exposure to conditions akin to operational settings during data acquisition.

2) Timeliness: IoT devices typically aim to capture real-time or near-real-time data. Thus, the data transmission, processing, and access mechanisms must ensure that data is accessible with minimal delay, adequately serving the application's temporal requirements.

3) Completeness: This aspect ensures that collected data points span the requisite time period necessary for deriving meaningful insights, safeguarding against gaps in data coverage.

By adhering to these principles, IoT devices can yield high-quality data essential for informing accurate and timely decision-making within the framework of digital twins.

Example: Dynamic Air Quality Data calibration process

In COMPAIR, data from citizen science sensors, professional sensors, and high-end official measuring stations were amalgamated to inform and evaluate local planning decisions. Hofman et al. (2020), COMPAIR deliverable D3.5 and D3.6 highlight the complexity of monitoring air quality in urban areas, emphasising the necessity for high spatial and temporal resolution to accurately assess population exposure. To address this, Hofman et al. propose an innovative IoT approach for highly granular air quality mapping in cities, relying on: i) a combination of cloud-calibrated fixed and mobile air quality sensors, and ii) machine learning methodologies to extrapolate spatiotemporal point measurements in both space and time. This approach was integrated into a calibration pipeline, utilised, and tested to enhance the data quality of citizen science sensors measuring particulate matter and black carbon (Figure 2).

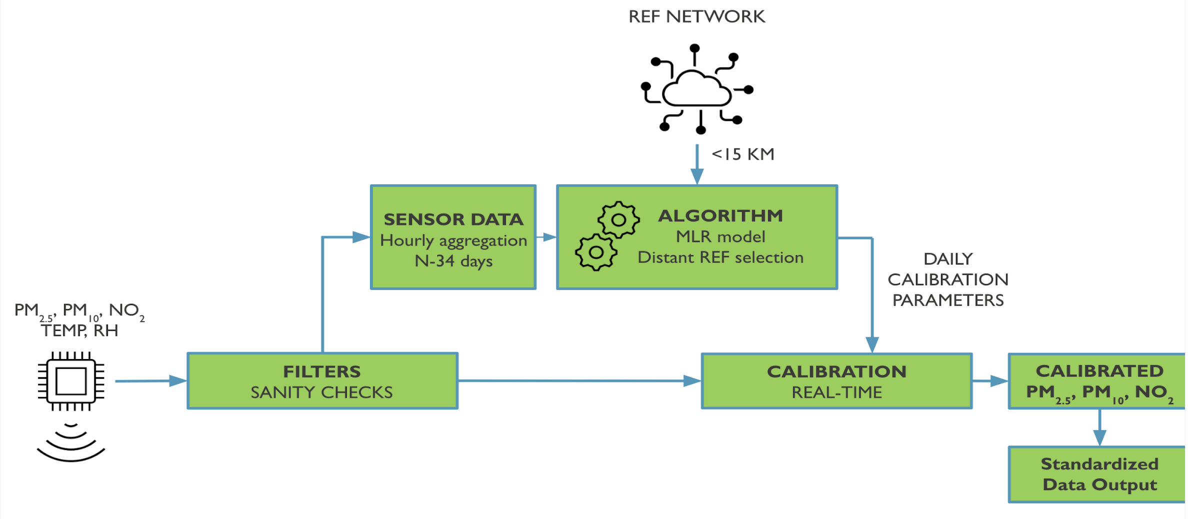

Figure 2: Cloud Calibration Pipeline for Citizen Science Air Quality Sensors

This pipeline integrates raw (potentially inaccurate) measurements from citizen science (CS) sensors with high-quality data from reference stations, enhancing the accuracy of measurements in real-time. The introduced approach offers a dual advantage. Firstly, cloud-based calibration eliminates the need for direct citizen involvement and manual intervention in addressing the intricacies of sensor accuracy. Secondly, scientific validation and harmonisation render the results readily accessible for integration into Local Digital Twins (LDTs), facilitating their utilisation.

IoT metadata and provenance

Metadata serves as a structured descriptor of data, facilitating its discovery by both systems and humans. Řezník et al. (2022) underscore metadata's pivotal role in data-driven decision-making, highlighting its critical link to the reliability of decision outcomes by ensuring the reliability of the underlying data. It assumes a fundamental role during data collection, delineating dataset characteristics, provenance, and quality. Prominent metadata standards for dataset description include the Dublin Core, Content Standard for Digital Geospatial Metadata by the FGDC, ISO 19115 Geographical Data - Metadata, and Data Catalog Vocabulary (DCAT), encompassing a spectrum from basic to comprehensive information, including data provenance.

Borgini (2023) delves into another metadata dimension, focusing on IoT device metadata independent of the dataset. Borgini defines metadata as detailed information concerning IoT measurements, encompassing attributes such as generation time, originating system or device, and format. Ideally, metadata furnishes a standardised dataset description, facilitating comprehension and categorisation by diverse systems, applications, and resources.

In the context of digital twins, IoT assumes multiple essential roles spanning both levels of metadata. Firstly, metadata addresses interoperability challenges by swiftly enabling devices and systems seeking to interact with IoT devices to identify and connect using appropriate communication protocols. Moreover, metadata informs other devices about the data exchange capabilities of IoT devices, enhancing connectivity efficiency and reducing latency. Thus, metadata contributes to mitigating the interoperability challenge, particularly at the device level, and standardising the data catalogue of a Local Digital Twin (LDT) across both fast-moving IoT-generated data and slower-moving data such as terrain or height models.

6. IoT Data Standards and Semantics

To grasp the significance of semantic interoperability in rendering IoT data, it is imperative to delve into the landscape of existing IoT data standards and their respective objectives. Thus, we embark on a comprehensive review encompassing several pivotal generic and geospatial IoT standards, setting the stage for a detailed examination of minimal interoperability requisites essential for sharing context information, data models, and transactions.

6.1. IoT data standards

Throughout the years, numerous IoT standards have emerged, spanning domains such as construction, agriculture, health, and smart cities. These standards encompass not only syntax and semantics but also crucial elements like security. This overview concentrates on IoT standards established by renowned international standards-setting bodies such as ISO (International Standardisation Organisation), IEEE (Institute of Electrical and Electronics Engineers), W3C (World Wide Web Consortium), and OGC (Open Geospatial Consortium). Additionally, readers keen on exploring further can reference the efforts undertaken by telecommunication entities such as ITU (International Telecommunication Union) and ETSI (European Telecommunication and Standardisation Institute).

GENERIC standards

ISO/IEC 30141 - Internet of Things Reference Architecture

ISO/IEC 30141 establishes a standardised IoT Reference Architecture incorporating a shared vocabulary, reusable designs, and industry best practices. Employing a top-down methodology, it initially collates the essential attributes of IoT, abstracts them into a generic IoT Conceptual Model, and subsequently formulates a high-level system-based reference. This reference is then dissected into four architecture views—functional, system, networking, and usage—providing diverse perspectives.

The reference architecture serves as a foundational framework for the development of context-specific IoT architectures and, consequently, actual systems. These contexts may vary but are required to encompass the business, regulatory, and technological domains, including industry verticals, technological prerequisites, and nation-specific requirements. Furthermore, the architecture delineates functional requirements such as data and device management, security, confidentiality, and privacy, alongside non-functional requirements like maintainability, reliability, usability, high availability, and scalability of systems.

ISO/IEC 30161-1 ED1: IoT Data Exchange Platform for IoT Services

ISO/IEC 30161-1 outlines prerequisites for an IoT data exchange platform catering to a range of services within various technology domains. These include middleware components facilitating the co-existence of IoT services with legacy ones, endpoint performance across communication networks interconnecting IoT and legacy services, IoT-specific functions for efficient service deployment, the framework and infrastructure of IoT service communication networks, and guidelines for IoT data exchange platform implementation.

ISO/IEC 30165: IoT Real-Time Framework

ISO/IEC 30165 presents a comprehensive framework for strategically deploying Real-Time Internet of Things (RT-IoT) systems, with the aim of addressing common challenges encountered in real-time system development. Emphasising the critical nature of real-time capability, the standard goes beyond generic descriptions, acknowledging that deviations from precise timing constraints could lead to significant repercussions, including potential harm to individuals. This underscores the importance of adhering to stringent timing requirements in RT-IoT system implementation.

ISO/IEC 21823 1-4: IoT Interoperability

The ISO/IEC 21823 1-4 standards on IoT Interoperability comprise four parts. Part one, the 'framework,' addresses concerns pertaining to interoperability in communication among entities within IoT systems. Part two, the 'transport interoperability standard,' delineates a framework and stipulates requirements for transport interoperability, facilitating the construction of IoT systems capable of seamless information exchange, peer-to-peer connectivity, and communication both within and between different IoT systems. Part three, the 'semantic interoperability standard,' centres on IoT semantic interoperability, enabling data exchange between IoT systems using understood data information models. Part four specifies IoT interoperability from a syntactic perspective.

ISO/IEC 27400 and 27402, 27402.2: Cybersecurity - IoT Security and Privacy - Guidelines

The ISO/IEC 27400 standard offers guidelines concerning risks, principles, and controls for ensuring the security and privacy of IoT solutions. Additionally, ISO/IEC 27402 outlines baseline requirements for devices, providing guidance on a standard set of information and communication technologies (ICT) requirements necessary for IoT devices to support security and privacy controls. A thorough risk assessment is crucial for developing a risk treatment plan that identifies the requisite features and countermeasures for IoT devices. Management of systems employing IoT devices hinges on the capabilities of these devices. Building upon these baselines, vertical markets (e.g., health, financial services, industrial, consumer electronics, and transportation) can establish additional requirements tailored to the anticipated use and risks of IoT devices in their applications.

Institute of Electrical and Electronics Engineers (IEEE) IoT Standards

The Institute of Electrical and Electronics Engineers (IEEE) is a professional organisation established in 1963, responsible for a plethora of peer-reviewed journals, tutorials, and standards produced by its standardisation committees. Notable IEEE standards include P1912, addressing Privacy and Security Architecture for Consumer Wireless Devices, defining a unified communication architecture for diverse wireless communication devices. P1451-99 Standard for Harmonization and Security of IoT leverages the XMPP protocol's advanced capabilities to provide authenticated identities, authorization, presence, lifecycle management, interoperable communication, IoT discovery, and global provisioning. P2413 defines an Architectural Framework for IoT, elucidating various IoT domains, domain abstractions, and commonalities among different IoT domains. IEEE 802.15.4-2015 concentrates on sensor connectivity, delineating a protocol and compatible interconnection for data communication devices utilising low data-rate, low-power, and low-complexity short-range radio frequency (RF) transmissions in wireless personal area networks (WPANs).

GEOSPATIAL-oriented Standards

The Open Geospatial Consortium (OGC) pioneered IoT standards as early as 2011, and today, it continues to focus on syntactically and semantically oriented standards and standardised services. Some OGC standards have been developed collaboratively with the W3C. The standards outlined below centre on IoT modelling, including SensorML, IoT observations and measurements (OMS), Sensor Observation Service (SOS), Sensor Planning Service (SPS), and Common Data and Services (SWE).

OGC Sensor Model Language (SensorML)

The primary objective of Sensor Model Language (SensorML) is to provide a robust and semantically linked method for defining processes and processing components associated with the measurement and post-measurement transformation of observations. This encompasses sensors, actuators, as well as computational processes applied pre- and post-measurement. The overarching goal is to facilitate interoperability, initially at the syntactic level, and subsequently at the semantic level (by employing ontologies and semantic mediation), thus enabling sensors and processes to be comprehended more effectively by machines, employed automatically in intricate workflows, and readily shared among intelligent sensor web nodes. SensorML was ratified by the OGC in 2019 and formally published in 2020.

OGC Observations, Measurements, and Samples (OMS)

The Observations, Measurements, and Samples (OMS) standard delineates a conceptual schema for observations, the features involved in the observation process, and the features involved in sampling during observations. This model supports the exchange of information pertaining to observation acts and their outcomes, both within and across various scientific and technical communities. The Sensor Things API (STA), Sensor Observation Service (SOS), Semantic Sensor Network (SSN), and Sensor Open Systems Architecture (SOSA) all incorporate ISO19156 as part of their core. Developed jointly by the Open Geospatial Consortium and ISO, OMS was published as ISO19156:2023.

OGC SensorThings API (STA)

The SensorThings API offers an open, geospatial-enabled, and unified mechanism for interconnecting IoT devices, data, and applications via the Web. At its core, the SensorThings API facilitates two primary functionalities, each managed by distinct components. These components consist of the Sensing part and the Tasking part. The Sensing part furnishes a standardised approach for managing and retrieving observations and metadata from diverse IoT sensor systems, akin to functionalities provided by the OGC SOS. Meanwhile, the Tasking part mirrors functionalities akin to those provided by the OGC SPS. Approved by the OGC in 2020, the standard was subsequently published in 2021.

OGC Sensor Observation Service (SOS)

The SOS standard is tailored to scenarios necessitating the interoperable management of sensor data. This standard delineates a web service interface enabling the querying of observations, sensor metadata, and representations of observed features. Additionally, the SOS standard outlines procedures for sensor registration and removal, along with operations for inserting new sensor observations. Recognized as an OGC standard since 2016, the SOS standard offers a robust foundation for managing sensor data in diverse contexts.

Sensor Planning Service (SPS)

The OpenGIS® Sensor Planning Service Interface Standard (SPS) defines interfaces for queries furnishing insights into sensor capabilities and tasking methodologies. Designed to support queries with multiple objectives, including feasibility assessment of sensor planning requests, submission and reservation/commitment of requests, inquiry into request status, updates or cancellations of requests, and retrieval of information regarding other OGC Web services facilitating data access for requested tasks. The most recent iteration of the SPS standard, version 2.0, dates back to 2011.

Common Data and Services (SWE)

The Sensor Web Enablement (SWE) Common Data Model Encoding Standard establishes foundational data models for exchanging sensor-related data among nodes within the OGC® Sensor Web Enablement (SWE) framework. These models enable applications and servers to structure, encode, and transmit sensor datasets in a self-descriptive and semantically enriched manner.

Example: The importance of OGC SensorThings API

In the frame of the COMPAIR project, collaborative efforts among multiple technical stakeholders proved indispensable for the successful integration of diverse systems into a cohesive solution. A primary challenge encountered in this endeavour pertains to achieving consensus on communication interfaces, given the inherent diversity among partner systems. Additionally, the inclusion of citizen participation introduces another layer of complexity, as citizens may possess access privileges to specific segments of the solution.

Standardisation emerges as a pivotal tool to address these challenges, with the OGC SensorThings API assuming a central role. Functioning as a standardised framework, it establishes a shared foundation characterised by precise interface definitions, thereby harmonising concepts and terminology across disparate systems. Moreover, the API mandates structured storage of sensor metadata and facilitates linkages to ontologies within measurement data. Consequently, both technical stakeholders and citizens alike benefit from enhanced comprehension of the data encapsulated within the solution.

W3C/OGC Semantic Sensor Network (SSN) ontology
The SSN ontology provides a formal framework for describing sensors, focusing on their capabilities, measurement processes, observations, and deployments. This ontology aims to facilitate semantic interoperability within physical sensor networks. Its primary concepts encompass sensors and their features, properties, observations, systems, measurement capabilities, operational and survival constraints, and deployment scenarios.

6.2. LDT-IoT-related initiatives

Digital twin-related standards are difficult to define due to the lack of clarity in defining what a digital twin is, on the one hand, and due to the system of systems approach, in which a large variety of applications, with specific standards, become part of a digital twin. The IoT-related digital twin standards almost all refer to DTs for industrial designs or digital twins in general, without specifying any particular domain. Nevertheless, the below generic standardisation initiatives related to digital twins are potentially useful for integrating IoT devices from the physical world into their LDT counterparts.

The ISO/IEC digital twin — Concepts and Terminology 30173:2023 initiative aims to define a common domain-overarching understanding. Wang et al. (2022) propose a five-dimensional digital twin model, with each dimension encompassing corresponding standards. The model differentiates between physical entities, virtual entities, data, connections, and services.

Physical entities within a digital twin system have according to Wang et al two primary functions: data collection and device control. These entities act as data sources and actuators for their virtual counterparts.

The IEEE proposed the IEEE 2888 standard series, which comprehensively defines the interface between the cyber world (digital twin) and the physical world. IEEE P2888.1 and IEEE P2888.2 establish the vocabulary, requirements, metrics, data formats, and APIs necessary for acquiring sensor information and commanding actuators, thereby defining the interfaces between the two realms. IEEE P2888.3 addresses the orchestration of digital synchronisation between the cyber and physical worlds. IEEE P2888.4 focuses on the architecture for a virtual reality disaster response training system, while IEEE P2888.5 and P2888.6 concentrate on the evaluation methods of virtual training systems and holographic visualisation for interfacing the cyber and physical worlds, respectively.

Virtual entities serve in a digital twin system as digital representations of physical entities. These virtual entities comprise models that describe physical entities across multiple temporal and spatial scales. The IEEE Standard on System Architecture of Digital Representation for Physical Objects in Factory Environments focuses on a non-smart-city digital twin environment and exemplifies how physical objects influence virtual entities in terms of objectives, components, data resources, and procedures.

The ISO TC184/SC4 and IEC/TC65/WG24 are developing standards to guide the implementation of an Asset Administration Shell (AAS), which aims to convert physical entities into digital twins. The AAS provides a semantic model comprising domain-specific sub-models that describe all information and functionalities of a given asset, including its features, characteristics, properties, statuses, parameters, measurement data, and capabilities (Dossogne, 2022). An asset registry is considered a fundamental feature of a local digital twin (LDT), requiring an index to catalogue all LDT assets. While asset documentation may exist external to an LDT instance, the registry is essential for traceability and discoverability. The Asset Registry offers a comprehensive overview of available datasets, schemas, vocabularies, algorithms, and other usable assets.

6.3. Minimal Interoperability Mechanisms for IoT

The Minimal Interoperability Mechanisms (MIMs) aim to unify cities and communities within a global market for solutions, services, and data based on their needs. To achieve these objectives, the MIMs define a set of practical capabilities based on open technical specifications to enhance the replicability and scalability of solutions (OASC, 2023a). The MIMs address a broad range of interoperability topics, many of which pertain to IoT-related issues. The most mature MIMs are MIM 1 (Context Information Management), MIM 2 (Shared Data Models), and MIM 3 (Ecosystem Transaction Management). The three above interoperability mechanisms are crucial for IoT interoperability and scalability, facilitating the use of IoT.

MIM 1: Context Information Management

Context information management ensures comprehensive and integrated access, use, sharing, and management of data across various solutions and purposes. It handles context information from Internet of Things (IoT) devices and other public and private data sources, offering cross-cutting context data and access through a uniform interface (OASC, 2023b). MIM 1 provides capabilities such as status information provision, access to multiple data sources, context information provision, information discovery and querying, and change and update management. An example of a MIM 1 implementation is NGSI-LD (ETSI-CIM, n.d-a).

MIM 2: Shared Data Models

The Minimal Interoperability Mechanisms (MIMs) for Shared Data Models include guidelines and a catalogue of common data models across various domains to enable interoperability among applications and systems in different cities. MIM 2 (OASC, 2023c) focuses on a harmonised representation of data formats and semantics, usable by applications that both consume and publish data. It employs the concept of Smart Data Models to ensure interoperability and replicability across multiple sectors, including all smart city domains. The primary capabilities of MIM 2 involve defining well-specified data models that capture the complete context, allowing applications to request specific attributes. MIM 2 recommends specifications such as NGSI-LD and SAREF (Daniele et al., 2020).

MIM 3: Ecosystem Transactions Management

The Ecosystem Transaction Management MIM 3 focuses on scaling data services within cities and communities, including those enabled by IoT and AI. MIM 3 necessitates easy and risk-free access to appropriate local data sources that are already available within these communities. A local data marketplace facilitates this access, allowing for the integration of relevant and available local data, solutions, and other resources, thus enabling the implementation of new and valuable services and solutions, many of which have been successfully deployed in other cities (OASC, 2023d).

The marketplace concepts of MIM 3 offer several capabilities: exposure of data and dataset offerings built on standard interoperability mechanisms, access to service offerings, and ecosystem transaction management. This includes effective matchmaking between relevant data sources (e.g., urban IoT data) from providers and respective data consumers, as well as trusted exploitation based on enforceable data usage agreements and secure value flow. A set of MIM 3 specifications is included in the SynchroniCity project (European Commission, 2022), the Marketplace Enablers Report (SynchroniCity, 2018a), and the Reference Architecture for IoT Enabled Smart Cities (SynchroniCity, 2018b).

7. Sustainable IoT Sensor Networks as part of a Digital Twin: Citizens Science

Čolaković and Hadžialić (2018) define Sensor Networks (SN) as a collection of sensors that communicate with each other and/or transmit data to other infrastructures, such as Fog or Cloud. An SN comprises sensors, actuators, firmware, and a thin layer of software framework. These components enable objects to be aware of their environment and to exchange data, aligning with one of the primary goals of the Internet of Things (IoT).

IoT systems are typically complex, primarily due to their influence on diverse aspects of human life and the deployment of various technologies that facilitate autonomous data exchange among embedded devices. The evolution of IoT significantly impacts multiple dimensions of human life, including security, safety, health, mobility, energy efficiency, and environmental sustainability. Therefore, it is crucial to address IoT-related issues and challenges from a holistic perspective, encompassing enabling technologies, services and applications, business models, and the social and environmental impacts. Many of these challenges are similar to those that LDTs aim to overcome as end user-oriented solutions.

This chapter will review two different types of sensor networks: professional and community-driven. Following this, we will present several use cases to demonstrate how these networks can be integrated within an LDT environment.

7.1. Professional sensor networks

Professional IoT sensor networks, as outlined below, offer a continuous stream of high-quality input data from the physical realm to the virtual domain. In an LDT context, multiple sensor networks can be integrated into a unified environment to support complex processes and use cases that benefit from collaboration among various sensor networks. Achieving syntactic and semantic interoperability at both the data and metadata levels is essential for effective joint knowledge creation across multiple sensors. Presented here are examples of professional IoT networks in urban settings with potential applications in LDT operational and policy contexts:

- Building management: Digital twins of buildings can integrate sensor data to monitor and optimise energy usage and ventilation based on occupancy and external conditions.

- Traffic management: Data from cameras, sensors, and IoT-connected vehicles can be used to simulate and optimise traffic flow, enhance road safety, and inform traffic planning.

- Environmental monitoring: Sensor data on water quality, flow, and levels can detect leaks, improve water distribution, and aid in flood and drought prevention. Air quality monitoring, combined with weather data, contributes to smog prevention and informs traffic and spatial planning.

- Waste management: Digital twins of waste management systems utilise sensor data to monitor bin levels and optimise waste collection routes for municipal waste authorities.

- Smart grids: An LDT of the electrical grid integrates real-time data from smart metres, power line sensors, and substations to monitor energy consumption, identify faults, and facilitate energy distribution.

In the ubiquitous deployment of IoT, managing the vast volume of generated data is a significant concern. Cities are inundated with data from various sources, often referred to as a "data tsunami." For instance, a single smart building can house thousands of sensors, exemplified by Edge in Amsterdam with 28,000 sensors. Scaling this to the city level involves numerous nodes, including buildings, lamp posts, waste bins, parking lots, and other urban infrastructure, posing integration challenges. Beyond integration, managing big data entails addressing storage, computing power, privacy, and environmental sustainability concerns associated with data centres and supercomputers processing incoming data streams.

Moreover, the hype surrounding big data has prompted questions regarding its relevance. Critics argue that for many problems, small data suffices. In the IoT context, small data refers to specific attributes of current states or measured conditions, such as location, temperature, vibration, and movement. These small datasets often trigger events and communication between physical assets. Small data packets can be transmitted over long distances using technologies like Narrowband IoT, LTE-M, and LTE-MTC, commonly found in devices used for citizen-led monitoring.

7.2. User-Generated Content and Citizen Science IoT Networks

The utilisation of Citizen Science IoT devices to bridge the gap between the physical realm and the digital world is a relatively recent development. One of the earliest and most notable instances of a do-it-yourself (DIY) IoT citizen science air quality sensor initiative was the grassroots Luftdaten (n.d.) project in Germany, initiated in 2015. Following the success of the Luftdaten initiative, the Sensor.Community (n.d.) platform emerged as a multi-measurement successor, boasting a user base of over 10,000 individuals worldwide. The CurieuzeNeuzen Garden (n.d.) project in Flanders, Belgium is a noteworthy large-scale endeavour exemplifying a quadruple helix IoT initiative. To monitor heat and drought levels, this initiative deployed NB-IoT soil sensors in approximately 5,000 locations, primarily private gardens selected from a pool of over 50,000 candidates. The project garnered widespread attention due to collaboration among research institutes, the Flemish government, newspaper media, private enterprises (including telecom providers), and prominent nature conservation organisations. The data's availability as open data, coupled with robust scientific backing—particularly evident in the CuriezeNeuzen Garden project—renders it appealing for integration into e.g. a Local Digital Twin (LDT) simulation and decision support framework.

7.3. Citizen Science IoT data for policy making: new possibilities and potential risks

When considering the integration of citizen science as a new data source comparable to government and private sector data, along with its potential to fuel policy instruments like e.g. LDTs and associated policy-making processes, several arguments and risks merit examination. Utilising an IoT network within the framework of citizen science presents both novel possibilities and associated risks. Drawing from experiences in the COMPAIR project across five pilot locations spanning Belgium, Bulgaria, Germany, and Greece, we summarise these considerations as follows:

Opportunities:

- Access to diverse locations: Citizen science offers access to a wide array of locations, including private spaces such as gardens and facades where individuals reside or work;
- High network density: The framework provides a high number of measurement spots at a reasonable cost, facilitating potential variability or spatial concentration;
- Lower operational costs: External management of sensors by volunteers enables efficient and continuous data collection at reduced operational expenses;
- Citizen involvement: Participants are actively engaged and exhibit a keen interest in measurement results and the subsequent policy decisions.

Risks:

- Lack of control over devices: There may be instances of devices being disconnected without oversight;
- Increased likelihood of failure: Challenges such as unstable network connections and malfunctioning sensors may arise;
- Higher probability of incorrect measurements: Factors like changes in location or measurement direction can lead to inaccuracies;
- Volunteer dropout: Volunteers may disengage from participation over time;
- Unequal spread: Challenges in network setup may occur due to lack of interest or installation options in certain areas, resulting in uneven coverage;
- Connectivity issues with IoT sensors: Some IoT sensors may experience connectivity challenges, impacting data collection.

An important insight from the COMPAIR project is the variation in network connectivity across the EU. While NB-IoT and LTE-M coverage is available in most European countries, comprehensive country-wide availability is not guaranteed, particularly in regions such as the Balkans. Discrepancies between urban and rural areas are also evident, with good connectivity in cities contrasting with limited coverage in rural settings. In cases where IoT infrastructure is insufficient, data collection becomes untenable, thereby affecting citizen participation. It is imperative for projects to conduct thorough connectivity assessments before determining what, where, and how to measure using which sensors.

8. IoT Connectivity - Lessons learned from COMPAIR - Need for a level-playing field

This IoT connectivity chapter is about the field experiences the COMPAIR project sensor providers faced during the COMPAIR project. These issues have been formulated as a note that has been discussed with the EC. The text below is based on this note and describes issues experienced, mitigation measures undertaken, and recommendations to create a level playing field for IoT sensor roll-out in the European Union.

8.1. Relevance

The COMPAIR project experienced difficulties connecting IoT devices suitable for Citizen Science and Smart City measurements via low-power narrowband communication systems. As a result, citizens in cities and regions lacking good connectivity were effectively excluded from participating in citizen science experiments.

During piloting work, the COMPAIR consortium noticed considerable connectivity differences between partners in various member states. Differences which hinder the future roll-out of citizen-driven smart cities all over Europe. The project believes actions at a European scale are essential to ensure narrowband connections are available alongside 5G and 6G connectivity, allowing IoT innovation across Europe. Narrowband networks in countries with an existing Citizen Science project culture are better available (e.g. Belgium, The Netherlands, and Germany) than in countries with a more limited tradition of rolling out IoT sensors for Citizen Science purposes (e.g.Greece and Bulgaria). If left unchecked, the current situation will further widen the citizen inclusion and participation gap between member states now that Citizen Science projects have shifted towards using IoT-based sensors.

8.2. IoT connectivity standards used during the COMPAIR project

COMPAIR stands for ‘Community Observation Measurement & Participation in AIR Science’
and is a citizen science project using air quality and traffic counting IoT sensors, specifically (among other sensors) SODAQ AIR[13] & Telraam S2[14]. These IoT sensors are built for (near) real-time monitoring and require constant connectivity to push data to the cloud. The COMPAIR sensors use the common IoT telecom connectivity standards LTE-M and NB-IoT, which are, in principle, readily available in all EU regions and cities.

- LTE-M[15] or LTE-MTC ("Long-Term Evolution Machine Type Communication") is a type of low-power wide-area network radio communication technology standard developed by 3GPP for machine-to-machine and Internet of Things (IoT) applications.
- Narrowband Internet of Things (NB-IoT) is a low-power wide-area network (LPWAN) radio technology standard developed by 3GPP for cellular network devices and services.

A third common IoT connectivity standard is LoRa. However, LoRa is not within the domain of the Telecommunications or Communication Service Providers (operating on so-called "free bands":

[13] https://sodaq.com/products/air/

[14] https://telraam.net/nl/S2

[15] Note that LTE-M is not to be confused with LTE (4G). The latter is the typical network for smartphone/cell phone connectivity and is omni-present. However, LTE is designed for high data throughput and is too expensive and power-hungry to be used for simple connected devices, especially in a citizen science context. For that reason, specific radio communication technology is developed and gradually implemented for many years for connected IoT devices: LTE-M & NB-IoT.

868MHz/EU and 915 MHz/USA) and typically depends on local communities or businesses to provide connectivity. LoRa has a disadvantage in that the connectivity depends on local communities providing connectivity through LoRa gateways locally installed at numerous locations, which is difficult to maintain. Another disadvantage is its low capacity and limited air time to an "on air time" of 1%.

8.3. Problems experienced

COMPAIR sensors were deployed in cities in Belgium, Germany, Greece and Bulgaria. IoT connectivity was seamlessly available in all countries except Bulgaria. Although the deployment of sensors occurred in 2 major Bulgarian cities (Sofia & Plovdiv), where we expect good coverage for at least one of the IoT radio communication technologies (LTE-M or NB-IoT), it was not possible to establish connectivity for any of the sensors.

Looking at the bigger picture, we see T-Mobile/Deutsche Telekom, one of the leading European telcos, does not provide LTE-M/NB-IoT connectivity in Bulgaria. The same holds for the UK-based Vodafone. These companies typically work with roaming agreements with national telecom providers. It is concerning that these companies do not provide LTE-M/NB-IoT coverage in a number of EU countries in multiple Balkan countries, including EU member states.

COMPAIR also discovered that the LTE-M/NB-IoT coverage in the countryside is problematic, too. This is problematic for the roll-out of future non-city area-wide citizen science initiatives. An excellent example of an impactful large-scale citizen science initiative rolled out using NB-IoT “CurieuzeNeuzen Garden” in the whole of Flanders (5000 IoT sensors involved), isn’t transferable to other areas confronted with the same drought problem in France, Greece, Portugal or Spain.

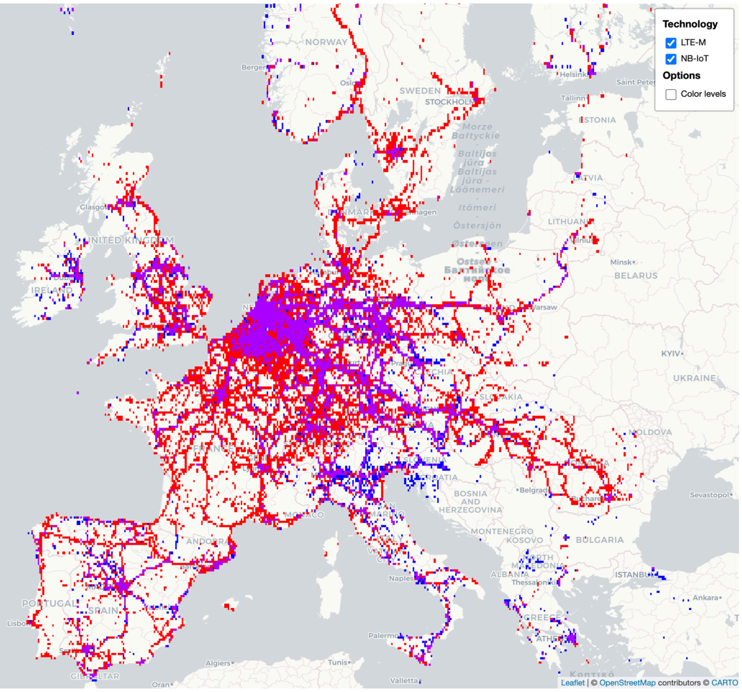

Figure 3: LTE-M and NB-IoT connectivity in Europe in November 2023 (Source: Sodaq)

8.4. Actions taken within COMPAIR

The COMPAIR consortium has undertaken several actions trying to resolve the issues. These actions, in sequence, were:

1. Contacted local telco operators to activate LTE-M/NB-IoT at least locally in specific areas of interest for the project, either directly or via intermediate service providers ⇒ this course of action had no impact whatsoever (no leverage over telco operators)
2. Changed operators (i.e. changed SIM cards) ⇒ this was not a solution, as all operators have the same issue. A1 is the only operator in BG claiming to offer NB-IoT coverage, but none of the sensor providers in COMPAIR could establish reliable NB-IoT connections due to unavailable coverage.
3. Explored retrofit solutions to convert connectivity to (classic) LTE ⇒ while this is technically a working solution, this action was considered not feasible as it involves a high cost (hardware change, expensive LTE modem, expensive data communication), which makes it unsuitable as a solution for citizen science IoT sensors, aiming for low cost and low power.

Insofar as possible, the COMPAIR consortium opted for a roll-back to wifi-connectivity, which posed a number of additional challenges on its own (cumbersome installation, the requirement of wifi connectivity nearby, the unreliability of Wifi, etc.). Wifi appeared to be an insufficient solution, very difficult to support and maintain, unsuitable for non-technical persons, and impossible to use when you want to work with vulnerable groups during Citizen Science projects.

8.5. Assessment and recommendations

The lack of NB-IoT AND LTE-M connectivity for connected devices in Bulgaria was an important limitation for implementing Citizen Science initiatives in the nation's capital and second-largest Bulgarian city. COMPAIR hadn’t expected to experience connectivity issues (at least in urban environments) while using the worldwide LTE-M and NB-IoT standards.

This assumption was based on public information from the operator's and intermediaries' own coverage maps (example 1 - example 2). The coverage maps do not reflect the reality on the ground, and operators claiming to offer NB-IoT & LTE-M are not providing this service.

For citizen science relying on IoT sensors data communication using LTE-M or NB-IoT, it is crucial to keep costs low. These network services are vital for citizen scientists and smart-city solution providers. Initiatives, like a regulatory framework and legislation, are urgently needed to ensure that Telco operators implement these technologies to ensure all EU citizens can access these connection services to create a level playing field for their involvement in shaping resilient, green and sustainable cities.

COMPAIR believes this course of action is possible, as the USA shows that a roll-out of NB-IoT and LTE-M by telecom operators across all its States is feasible.

Creating a European regulatory framework to ensure telecom operators provide LTE-M and NB-IoT is highly recommended. This action will ensure the future roll-out of low-cost IoT devices across Europe to allow everyone to support smart-city and citizen science initiatives, no matter which country they are based in.

Keeping connectivity and data exchange costs low is essential for equitable access to affordable IoT sensor networks, enabling citizens, including important groups like students, to measure and understand local air quality, mobility, climate impact, and more. The right to shape local policy and drive effective change in support of EU sustainability initiatives like the Green Deal 1 should not depend on which country you live in.

9. Conclusion

IoT networks serve as an essential building block for converting physical conditions into the digital representation of a city. Emerging network technologies and an expanding array of increasingly smaller and more affordable IoT devices are transforming how cities are monitored and managed. This white paper provides an overview of the standardisation needs and options for deploying integrated IoT networks providing data to data-dashboard and Digital Twin contexts, emphasising the necessity and applicability of Minimal Interoperability Mechanisms (MIMs). These MIMs are crucial for ensuring interoperability, which is necessary for the vision of a fully connected smart city to achieve. By examining concrete implementation examples, readers can learn how to address variations in data quality and leverage opportunities for stakeholder engagement, policy evaluation, and behavioural change facilitated by recent advances in IoT, citizen science, augmented reality (AR), data analytics, and digital twins.

During the COMPAIR project, the European Commission encountered and noticed structural problems related to the roll-out of IoT devices throughout the European Union. More specifically, the lack of NB-IoT AND LTE-M connectivity for connected devices in Bulgaria was a significant limitation for implementing Citizen Science initiatives in the nation's capital and second-largest Bulgarian city. COMPAIR had not expected to experience connectivity issues (at least in urban environments) while using the worldwide LTE-M and NB-IoT standards. The COMPAIR consortium advocate, since the technology itself is available, to force the telecom players to offer EU-wide low bandwidth and low power device support to create a level-playing field for IoT roll-out in and beyond a Citizen Science context.

Ensuring connectivity and data exchange remain simple and low cost is not just a matter of convenience, but a crucial element for providing equitable access to affordable IoT sensor networks. This accessibility empowers citizens, including key groups like students, to measure and comprehend local air quality, mobility, climate impact, and more.

10. References

Bonney, R. (1996). Citizen science: A lab tradition. *Living Bird*, *15*, 7–15.

Borgini, J. (2023, April 14). Use metadata for IoT data organization. *TechTarget*. https://www.techtarget.com/iotagenda/tip/Use-metadata-for-IoT-data-organization

Carfantan, G., Daniel, F., d'Orazio, L., Le, T.-D., Marin, X., Peau, O., & Rannou, H. (2020). Think Cities: The accelerator for sustainable planning. *2020 IEEE 36th International Conference on Data Engineering Workshops (ICDEW)*, 64–70.

Čolaković, A., & Hadžialić, M. (2018). Internet of Things (IoT): A review of enabling technologies, challenges, and open research issues. *Computer Networks*, *144*, 17–39. https://doi.org/10.1016/j.comnet.2018.07.017

Curieuze Neuzen. (n.d.). *Meer dan 50.000 CurieuzeNeuzen schreven zich in om de hitte en droogte te meten*. https://sensor.community/en/

Daniele, L., Garcia-Castro, R., Lefrançois, M., & Poveda-Villalon, M. (2020). *SAREF: the Smart Applications REFerence ontology*. ETSI. https://saref.etsi.org/core/v3.1.1/

Dossogne, V. (2022, November). A look at the latest standardisation projects in the field of digital twins. *S Innovation Forward*. https://www.sirris.be/en/inspiration/look-latest-standardisation-projects-field-digital-twins

ETSI-CIM. (n.d. a). *Industry specification group (ISG) cross cutting context information management (CIM)*. ETSI. https://www.etsi.org/committee/cim

European Commission. (2022, August). *CORDIS SynchroniCity: Delivering an IoT enabled Digital Single Market for Europe and Beyond*. https://cordis.europa.eu/project/id/732240

Hofman, J., Nikolaou, M. E., Huu Do, T., Qin, X., Rodrigo, E., Philips, W., Deligiannis, N., & La Manna, V. P. (2020). Mapping Air Quality in IoT Cities: Cloud Calibration and Air Quality Inference of Sensor Data. *2020 IEEE SENSORS*, 1–4. https://doi.org/10.1109/SENSORS47125.2020.9278941

Irwin, A. (1995). *Citizen science: A study of people, expertise, and sustainable development* (1. publ). Routledge.

Kerson, R. (1989). Lab for the Environment. *MIT Technology Review*, *92*(1), 11–12.

Luftdaten. (n.d.). *Luftdaten, Measure Air Yourself*. https://luftdaten.info/

Miller-Rushing, A., Primack, R., & Bonney, R. (2012). The history of public participation in ecological research. *Frontiers in Ecology and the Environment*, *10*(6), 285–290. https://doi.org/10.1890/110278

Newman, Greg, Andrea Wiggins, Alycia Crall, Eric Graham, Sarah Newman, and Kevin Crowston. "The Future of Citizen Science: Emerging Technologies and Shifting Paradigms." *Frontiers in Ecology and the Environment* 10, no. 6 (2012): 298–304. https://doi.org/10.1890/110294

OASC. (2023a). *Minimal Interoperability Mechanisms – MIMs*. https://oascities.org/minimal-interoperability-mechanisms/

OASC. (2023b). *OASC MIM1: Context Information Management*. OASC. https://mims.oascities.org/mims/oasc-mim-1-context

OASC. (2023c). *OASC MIM2: Shared Data models*. OASC. https://mims.oascities.org/mims/oasc-mim-2-data-models

OASC. (2023d). *OASC MIM3: Ecosystem Transactions Management*. OASC. https://mims.oascities.org/mims/oasc-mim-3-contracts

Open Knowledge Belgium. (2018). *InfluencAir—Citizens measuring air quality*. SlideShare. https://www.slideshare.net/slideshow/influencair-citizens-measuring-air-quality/89641201

Pawar, P., & Trivedi, A. (2019). Device-to-Device Communication Based IoT System: Benefits and Challenges. *IETE Technical Review*, *36*(4), 362–374. https://doi.org/10.1080/02564602.2018.1476191

Řezník, T., Raes, L., Stott, A., De Lathouwer, B., Perego, A., Charvát, K., & Kafka, Š. (2022). Improving the documentation and findability of data services and repositories: A review of (meta)data management approaches. *Computers & Geosciences*, *169*, 105194. https://doi.org/10.1016/j.cageo.2022.105194

Rp, J., K, R., A, A., & K, L. N. (2021). IoT in smart cities: A contemporary survey. *Global Transitions Proceedings*, *2*(2), 187–193. https://doi.org/10.1016/j.gltp.2021.08.069

Stock, O. (Ed.). (1997). Spatial and temporal reasoning. Kluwer Academic Publishers.

Syed, A. S., Sierra-Sosa, D., Kumar, A., & Elmaghraby, A. (2021). IoT in Smart Cities: A Survey of Technologies, Practices and Challenges. *Smart Cities*, *4*(2), 429–475. https://doi.org/10.3390/smartcities4020024

Synchronicity Consortium. (2018a). *D2.4. Basic data market place enablers*. Synchronicity Consortium. https://oascities.org/wp-content/uploads/2022/08/SynchroniCity_D2.4.pdf

Synchronicity Consortium. (2018b). *D2.10. Reference Architecture for IoT Enabled Smart Cities, Update*. Synchronicity Consortium. https://oascities.org/wp-content/uploads/2022/08/SynchroniCity_D2.10.pdf

Terziyski, A., Tenev, S., Jeliazkov, V., Jeliazkova, N., & Kochev, N. (2020). Meter. Ac: Live open access atmospheric monitoring data for Bulgaria with high spatiotemporal resolution. *Data*, *5*(2), 36. https://doi.org/10.3390/data5020036

Vohland, K., Land-Zandstra, A., Ceccaroni, L., Lemmens, R., Perelló, J., Ponti, M., Samson, R., & Wagenknecht, K. (Eds.). (2021). The Science of Citizen Science. Springer International Publishing. https://doi.org/10.1007/978-3-030-58278-4

Wang, K., Wang, Y., Li, Y., Fan, X., Xiao, S., & Hu, L. (2022). A review of the technology standards for enabling digital twin [version 2; peer review: 2 approved]. *Digital Twin*, *2*(4), Article 4. https://doi.org/10.12688/digitaltwin.17549.2