Data Integration in a Smart City: A Real Case

L. Davoli, L. Belli, A. Dall'Olio, F. Di Nocera, P. Adorni,
A. Cantelli and G. Ferrari

Abstract

The introduction and continuous integration of Internet of Things (IoT)-oriented technologies in urban environments leads to enhanced solutions in several domains (such as mobility, health, energy management, environmental monitoring, etc.), thus making a city "smart" and ultimately benefiting the everyday life of its citizens. As IoT systems are widely known to be producers of (often a very large amount of) heterogeneous data, in this chapter we discuss a modular and scalable approach to handle IoT-based data collection and management in a real smart city case, namely, that of the city of Parma, Italy. The proposed IoT infrastructure, the core component of which is a logical processing entity, acting as *middleware* and denoted as "city2i®," in charge of "digesting" the heterogeneous information generated by multiple data sources, allows the municipality to monitor the city status (from multiple perspectives) and to highlight "hidden" correlations among collected data.

1 Introduction

In recent years, the concept of *smart city* has been rapidly spreading, in connection with numerous urban contexts and particularly with the application of Information and Communication Technologies (ICTs). This is enabled by the cooperation of academic and industrial worlds, together with governments, institutions and citizens. From a general perspective, a smart city is a city that aims at becoming more livable, sustainable, and attractive for both citizens and potential visitors [1]. This goal requires the municipality to consider the involvement of different entities (such as businesses, multi-utility companies, public transport, and other service-related

---

L. Davoli · L. Belli · G. Ferrari (✉)
things2i S.r.l., Parma, Italy, and Internet of Things (IoT) Lab, Department of Engineering and Architecture, University of Parma, Parma, Italy
e-mail: gianluigi.ferrari@things2i.com

L. Davoli
e-mail: luca.davoli@things2i.com

L. Belli
e-mail: laura.belli@things2i.com

A. Dall'Olio · F. Di Nocera · P. Adorni
BT Enìa Telecomunicazioni S.p.A., Parma, Italy
e-mail: alessandro.dallolio@bt.com

F. Di Nocera
e-mail: francesco.dinocera@btenia.it

P. Adorni
e-mail: paolo.adorni@btenia.it

A. Cantelli
Comune di Parma, Parma, Italy
e-mail: a.cantelli@comune.parma.it

R. Menozzi (ed.), *Information and Communications Technologies for Smart Cities and Societies*,
The City Project 5, https://doi.org/10.1007/978-3-031-39446-1_2

stakeholders), which need to cooperate by creating working groups with heterogeneous competences from different city-related areas of intervention. Illustrative activities, which may increase the “smartness” of a municipality, include: efficient energy and resource management, environmental monitoring and protection, optimization of private and public vehicle usage and transport costs (e.g., targeting a shared mobility model, electric mobility, and car pooling), and others [2–5].

Besides, Internet of Things (IoT)-oriented infrastructures and technologies have significantly emerged in the last few years, supporting a large variety of devices, also denoted as *smart objects*, which can be helpful in the city context. For this reason, IoT technologies are considered as key enablers for the actual implementation of the smart city concept [6, 7], allowing the collection of huge amounts of data, which can then be synergistically managed through Artificial Intelligence (AI) algorithms and, in general, (big) data analytics [8]. Therefore, IoT deployment is spreading in different urban areas, with continuous growth of the number of heterogeneous connected devices equipped with different types of sensors and actuators. Considering these scenarios, it is clear that the communication infrastructure is the first key aspect to deal with in future smart cities, since connectivity is generally required *anytime* (during both daytime and night-time urban activities), *anywhere* (namely: in indoor environments, such as buildings; in outdoor environments, such us roads, parking areas, and parks; and also with mobile nodes, such as vehicles or bikes), and *anyhow* (between wireless personal devices —ranging from citizens’ laptops, smartphones, and tablets— to constrained and battery-powered IoT devices).

IoT smart objects deployed in a city are often characterized by different communication technologies and capabilities, and interact with the environment according to different paradigms, e.g., Human-to-Machine (H2M) or Machine-to-Machine (M2M). Taking into account different urban application scenarios, a heterogeneous set of communication technologies and different network architectures may be adopted, aiming at guaranteeing reliability and efficiency depending on the specific features of the services to be provided to citizens. In most cases, connectivity among IoT devices is guaranteed through flexible and low-cost wireless connections, e.g., through wireless mesh networks [9–11]. Generally, wireless short-range communication technologies are preferable when *end-to-end* communications are not possible because of power limitations or physical obstacles. In these cases, *multi-hop* device-to-device communications can be employed to transfer information from IoT nodes (the data sources) to data consumers (e.g., border routers, Cloud servers, users’ applications, etc.). In other deployment conditions, *star* network topologies, based on long-range communication technologies and enabling *one-hop* communications, are preferable. This can be the case of outdoor urban environments where direct and reliable communication links between a central “hub” and several IoT nodes within its coverage are available. Relevant examples of viable solutions are IoT deployments based on Long Range Wide Area Networks (LoRaWAN) [12] and NarrowBand IoT (NB-IoT) [13], currently attracting much interest in smart city scenarios where a large coverage is needed, with minimum infrastructure support and energy consumption.

The complexity of urban scenarios prevents the use of a single connectivity technology for all the possible monitored areas and, consequently, for all the services that can be implemented. Rather, a hybrid and flexible networking infrastructure, combining heterogeneous technologies, is required. The most relevant communication technologies employed in smart city contexts can be summarized as follows, on the basis of the transmission range and communication data rate [14–18], with pros and cons detailed in Table 1.

- Low-Power Wide Area Networks (LPWANs): these networks are designed to interconnect battery-powered devices with very low

Table 1 Pros and cons of relevant communication technologies employed in smart city contexts

Technology	Pros	Cons
LTE-M	Good coverage and roaming, acceptable power consumption, suitable for mobile and static devices	Not adapt for large amounts of data, target low-rate applications, asymmetric data transfer (with limited downlink channels)
NB-IoT	Good power consumption in poor coverage conditions	Suitable only for static devices
LoRa/LoRaWAN	Low power consumption and cost, secure bidirectional communication	Low data rate and throughput
4G/5G	High data rate, bandwidth, coverage, and flexibility	Additional SIMs required (until eSIMs will not be natively integrated into equipments), M2M traffic to be accommodated
IEEE 802.11	Lack of wires, user can move, no need to be stuck at one place	High signal attenuation, limited service radius, less stable compared to wired connections
IEEE 802.15.4	Power saving, collision avoidance, low cost, open standard	Short range, low data rate and throughput, lack full interoperability
BLE	Low cost, easy to install, native support in modern equipments	Short-range communication, secure flaws

bit rate (on the order of b/s) over very long distances (on the order of km). Due to their low cost, wide coverage, and straightforward setup, LPWANs are being deployed in scenarios where small amounts of data have to be transmitted by a large number of devices [19]. LPWANs typically operate in both unlicensed and licensed frequency bands, and are based on different (open or proprietary) standards, such as LTE-M [20] and the above-mentioned NB-IoT[1] and LoRaWAN.

- Cellular (e.g., 4G/5G): thanks to its improved features, such as very low latency (below 1 ms) and very high bandwidth (over 10 Gb/s), 5G is emerging as one of the primary enablers for IoT. Hence, 5G can be considered as an enabling communication technology for smart cities, allowing thousands of IoT devices to be connected, regardless of their location, and supporting applications such as smart traffic systems, public safety, enhanced tourism [21], security and surveillance [22].
- Wireless Local Area Networks (WLANs) and short-range networks: many use cases in smart cities require the deployment of "regional" (namely, covering a limited spatial area) and, in some cases, "individual" networks (such as Personal Area Networks, PANs). In this context, available communication technologies span from IEEE 802.11 to IEEE 802.15.4 and Bluetooth Low Energy (BLE) [23].

A heterogeneous communication architecture allows continuous and ubiquitous sensing in urban contexts, leading to large volume of produced data, which can be represented according to multiple formats. These data potentially represent a valuable asset for cities since, if properly processed, they can help to significantly

---

[1] LTE-M and NB-IoT are formally a cellular technology. However, the use of the available resources is typical of an LPWAN and, therefore, we consider these technologies in this category.

improve the efficiency and effectiveness of several operations related to city management and administration. However, because of the absence of common platforms and clearly-defined (and open) standards, the majority of data produced and collected in a city are generally stored and secured in proprietary data centers, only accessible with dedicated software solutions that act as isolated “verticals” developed by the vendors.

Recently, the development and deployment of frameworks or middlewares targeting smart cities has been addressed in the academic literature. As a reference, [24] presents an overview of the wide landscape of IoT-enabled smart cities, considering a selection of relevant municipal initiatives around the world having IoT as the key enabler for realising the smart city paradigm. More in detail, the objective in [24] is to identify the key IoT technologies concurring in the smart city realization and the key challenges that are currently being addressed for the success of smart cities. Among other aspects, the following problems are highlighted: (i) solutions based on the use of proprietary IoT platforms cannot be replicated across cities that often suffer from vendor lock-in (i.e., they cannot easily buy and integrate components or applications from different vendors); and (ii) public administrations cannot, on their own, fully take advantage of the smart city paradigm since they are consumers of services, but they are not actively engaged in the smart city as producers of valuable information that would bring forward new services, and contribute to a virtuous cycle with many benefits.

The problem of interoperability between different smart city systems has been addressed also in [25], where a set of requirements is defined for interoperability (in accordance to five different aspects for the smart city concept, namely, syntactic, semantic, network, middleware, and security) after collecting and analyzing studies on heterogeneous IoT systems.

Moreover, additional surveys have been presented in the literature in the last few years [3, 26–28], all intended to clarify and highlight possible ways in which IoT technologies can be employed in infrastructure projects to enhance both productivity and responsiveness. In [26, 28], projects or frameworks are presented focusing on the following specific topics or services: transportation, energy, parking, waste management, general monitoring, and healthcare. Other works do not focus on a single specific application or topic, but provide a more general approach [29, 30]. [27] presents IoT frameworks for smart city applications, in turn comparing technologies and architectures, and finding that they share, in general, an architecture with the following four layers: (i) a sensing layer, (ii) a network layer, (iii) a middleware layer, and (iv) an application layer.

Based on these survey results, an abstract five-layered IoT framework concept for smart city applications is proposed to address the needs of smart city applications as a reference for successful IoT framework implementations. To this regard, the main aspects addressed by the proposed framework concept are sustainability, decentralization, autonomy, security, and modularity.

With respect to the aforementioned state of the art, the main novelty of the middleware-oriented architecture proposed in this chapter is related to the actual implementation of a framework targeting the municipality of a smart city (since citizens are often targeted by applications provided by vendor’s commercial platforms); the goal is to integrate data collected by different systems, thus building a new tool for accurate understanding of city processes.

It is also important to remind that independent and isolated “verticals” developed by vendors: (i) are not natively designed to interoperate and communicate with each other, and (ii) do not aim at contributing to a common urban data management infrastructure. This motivates an always-increasing need for novel, open, and scalable platforms and frameworks that can help in the collection, integration and analysis of complex smart city data. This is fundamental to extend the services that a “smart” city can provide to its administration, citizens and visitors, thus building a new data management and utilization paradigm.

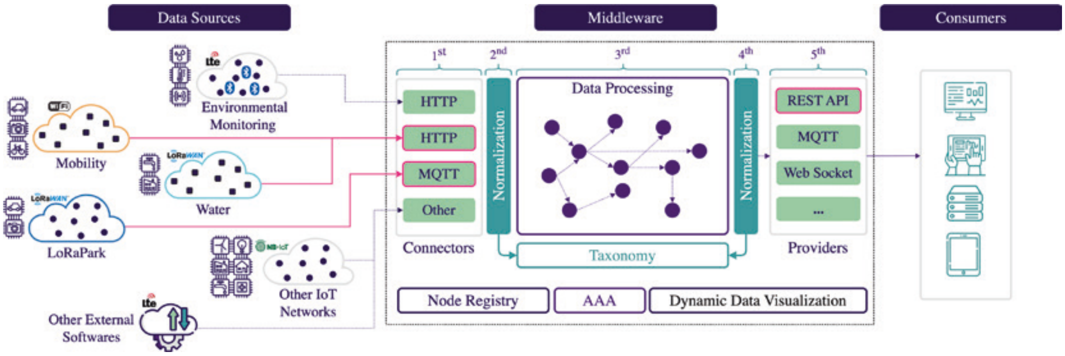


Fig. 1 Overall representation of the proposed IoT-oriented infrastructure working on top of its core component, denoted as middleware. The components and modules currently defined and deployed in a Proof of Concept (PoC) in the city of Parma, Italy, are highlighted with pink solid lines and colored clouds.

Motivated by the previous observations, in this chapter we present and discuss a novel, modular and scalable approach to handle IoT-based data collection and integration in a real smart city case, given by the city of Parma, Italy. The proposed IoT infrastructure, the core component of which is a logical processing entity acting as middleware and denoted as “city2i®,” is designed to extract shareable knowledge from heterogeneous smart city data. The middleware, on top of a heterogeneous data collection infrastructure, allows to transfer streams of data from IoT devices (Far Edge) to the Edge and, eventually, to the Cloud. This, in turn, provides the municipality and the citizens with a tool that, by integrating different systems, allows to monitor the entire city status (from multiple perspectives), and also to highlight “hidden” correlations among (IoT) data potentially collected by different (IoT) “verticals.”

2 A New Data Integration Environment

2.1 The Middleware Architecture

The middleware is a modular software platform representing the core of the proposed smart city-oriented infrastructure. Its main goal is to integrate the data generated by all the possible data sources that are available in the city context, such as heterogeneous IoT networks, external software services used by the municipality administrators, public open datasets available in the Cloud, and so on. The high-level architectural representation of the middleware, together with its connections to external entities (data sources and consumers at the input and at the output, respectively), is shown in Fig. 1.

The data sources may be highly heterogeneous and correspond to IoT deployments and software services (with different purposes) active in the urban context. IoT data sources rely on specific technologies (e.g., sensing resources), communication protocols (e.g., LoRaWAN, WiFi, BLE, TCP or UDP, MQTT, CoAP [31], etc.) and data formats, depending on the tasks to be performed and the required performance. As an example, Fig. 1 includes a LoRaWAN-based IoT network deployed for the city environmental monitoring, including temperature, humidity, and pollution sensors. Another scenario may involve a WiFi-based IoT deployment to monitor mobility flows in specific areas of the city (e.g., pedestrian, bike or vehicle counters on the roads). Other possible data sources may involve external software acquired by the municipality (e.g., administrative software), as well as *open data* and services deployed in the Cloud: in this case as well, data sources may rely on a specific approach to transfer information (e.g., REST APIs, WebSocket, MQTT, etc.) and data formats (e.g., JSON).

The main components of the proposed middleware-oriented platform can be summarized considering the following five layers highlighted in Fig. 1:

1. *Connectors*, needed to integrate data sources;
2. *Normalization module*, responsible for knowledge extraction;
3. *Taxonomy module*, required to represent smart city information;
4. *Data Processing module*, in charge of custom processing on the data;
5. *Normalization module*, responsible for representing new data.

In the following, all these components will be further described, together with other additional modules required for the proposed architecture's life cycle.

Each data source is integrated in the middleware through specific software components, denoted as *Connectors*, which are implemented to integrate data streams from a specific (and, from a top-down perspective, *atomic*) data source. The set of connectors can thus be considered as the *first layer* of the middleware. This layer has to be extended, with the implementation of a new connector, to support the integration of any new data source (e.g., if the municipality decides to deploy a new IoT system in the city). In fact, connectors depend on the specific technology and paradigm implemented by the manufacturer/vendor of the newly deployed system.

The *second layer* of the middleware corresponds to a component, denoted as *Normalization module* (which is also present in the output layer), responsible for knowledge extraction. This module, in accordance with the platform's *Taxonomy module* governing the information representation rules and policies valid inside the middleware itself, translates *raw* data acquired by connectors into one or more integrated information streams, with a common data structure that can thus be managed by subsequent middleware modules. More in detail, the goal of the Normalization module is to create high-level, geo-localized, and integrated *data overlays* identifying and semantically labeling input data streams (e.g., temperature values, parking information, air pollution indicators, etc.) that may be obtained by completely different data sources, but that are logically coherent. In practice, the Normalization module maps each *raw* piece of information generated by data sources to one or more taxonomy's entries (on the basis of its ontological meaning) and associates each value with an absolute timestamp and geographic coordinates. The result of this preliminary operation is a set of basic data overlays, composed by similar data, obtained through the integration of different data sources by the middleware. As an example, a "city temperature overlay" may derive from the integration of geo-localized heterogeneous temperature sensors (developed by different manufacturers and vendors) deployed in the city with open data publicly available in the Cloud. This kind of data integration is instrumental to subsequently manage information in a more contextualized way, regardless of the specific technology used by sensors to collect data, and by data sources to transmit them.

The *third layer*, corresponding to the *Data Processing module*, has been designed to perform processing tasks on basic "normalized" data overlays generated by the previous layer. The structure of this module is intended to be highly modular and configurable, following a Big Stream-oriented approach [32, 33] able to cope with very large amounts of information to be managed. The Data Processing module is composed by a set of modular processing units that: (i) take one or more data streams as input; (ii) perform custom operations defining the *behavior* of each "atomic" processing node; and (iii) generate one or more information streams, related to one or more taxonomy's entries, as output. The operations performed on the data can span from simple statistical tasks (e.g., mean, standard deviation, etc.) to complex data fusion tasks, as well as AI-based processing. To this regard, pictorial examples of how the AI can intervene in the data processing may include: (i) air quality prediction based on pollutant levels sensed by IoT devices spread across the city; (ii) traffic and mobility flow management based on cellular signals' strength obtained from *third-party* data sources; and (iii) water consumption forecasting in public buildings based on data collected through smart metering solutions. There is no actual limit to the specific operations that can be implemented in single processing units, as the set may be freely extended depending on the final data overlays of interest for smart city administrators, as well as for citizens and visitors. The aforementioned processing units can be defined and included in the *Data Processing module* as needed by the smart city municipality, building a sort of graph where data streams connect different processing steps. We remark that this highly-modular infrastructure introduces latency at each processing unit "crossed" by multiple data streams. These latency components are, in general, unpredictable, depend on the specific algorithm or computational complexity needed for the implementation of the desired behavior, and are independent of the architecture itself. Nevertheless, the middleware has been designed to minimize the latency needed to forward data between two consecutive processing units [32, 33].

The *fourth layer* is associated with a *Normalization module* similar to that of the *second layer* and dedicated to data provisioning. Hence, the information output by this module is normalized cooperating with the internal cross-layer Taxonomy module, in order to provide data with a format coherent with the taxonomy used in the overall system and with the reference domain of the smart city.

Finally, the generated high-level information should be output by the middleware to interested external entities: this can be performed through different sets of data access mechanisms, which are denoted as *providers* and constitute the *fifth layer* of the middleware. In detail, each provider may act in a *passive way*, waiting for requests from external entities, as well as in a *proactive way*, publishing data as soon as these are available. Relevant examples of data provisioning are HTTPS REST APIs, subscriptions to MQTT topics, and WSs. Different providers' modules can be implemented to extend the platform and to satisfy the requirements given by external applications or developers, interested in creating innovative services using the information overlays generated by the proposed middleware.

The following additional modules, with respect to those of the five operational layers, are relevant in the life cycle of the proposed IoT-oriented architecture.

- *Node Registry module*: in charge of maintaining the list of IoT nodes acting as data sources for the IoT infrastructure, with information related to the nodes attached to data streams. This information may be useful for data consumers.
- *Authentication, Authorization, and Accounting (AAA) module*: responsible for controlling the access to the resources by both connectors and providers.
- *Dynamic Data Visualization module*: in charge of visually representing, upon request, the information handled by the IoT middleware. Being part of a modular solution and aiming to adhere to the middleware's taxonomy, the design of the Graphical User Interface (GUI) should reflect the heterogeneity of the IoT data flows, too [34]. In particular, it should allow to dynamically build customized User Interfaces (UIs) on the basis of data and information descriptors (e.g., XML- or JSON-based).

The aforementioned modules allow the proposed IoT architecture to provide smart cities' administrators with a unified tool, that can be seen as a sort of city's control panel. This allows to extract knowledge in a "simple" way from heterogeneous urban data sources, hiding technical aspects to the municipality, citizens and visitors, since they lack the cecessary technical skills. As an example, heterogeneous data streams can be integrated and classified through the taxonomy module as "similar," even if the physical data sources are completely different (e.g., manufactured by different vendors). This allows to have a comprehensive overview on the available data from the entire city perspective, which on the other hand is generally prevented by limited vertical applications provided by the single systems' vendors.

As detailed before, the cross-layer Taxonomy module is crucial in the proposed middleware architecture. In fact, it provides a unified ontology to label the information overlays inside the proposed infrastructure. This module allows to classify and normalize data provided by both external data sources (after the acquisition step through the connectors) and the middleware itself (in the Data Processing module), providing a coherent information flow to external customers. The resulting classification is then shared to interested entities through providers. To this end, the entries composing the taxonomy should be selected in order to be assigned to possible data generated by IoT objects deployed in the smart city, such as, for example: environmental monitoring data (e.g., indoor and outdoor temperatures, pressure, humidity, air- and weather-related parameters, etc.), energy and resources utilization data, mobility- and traffic flow-related data, and others. As mentioned before, besides IoT systems, data sources can also be external and of a different nature (e.g., Cloud data): the taxonomy has to take this into account.

3 A Real Use Case: The City of Parma

In order to validate the proposed IoT-oriented middleware-based architecture directly *in the field*, inn this section we present an overview of its adoption (currently in progress) in the city of Parma, Italy. The city of Parma has implemented relevant Proofs of Concept (PoCs) to investigate how to offer enhanced services to its citizens and visitors, for example through the deployment of various LoRaWAN-based IoT systems, connected to the LoRaWAN city network infrastructure deployed by BT Enìa Telecomunicazioni S.p.A.[2]

One of the PoCs focuses on smart parking and urban mobility solutions.[3] More in detail, monitoring the parking availability is a relevant activity in a smart city context, as it is important both for citizens, who would like to find vacant parking spaces for their vehicles during their working days, and for the municipality, to verify parking sustainability (in terms of available parking spaces) and to detect improper utilization (e.g., missing payments, unlawful utilization of parking lots reserved for disabled people or for loading/unloading goods). Several IoT options are available for parking monitoring, depending on cost, number of lots to be monitored, and environmental conditions. In the city of Parma, the following two types of parking monitoring system have been evaluated in different downtown areas and are integrated as IoT data sources in the middleware.

- *Magnetic* IoT nodes (Parking Spot Sensor provided by Kiunsys[4]). A single IoT node is inserted directly in the asphalt in each single parking lot to be monitored, providing a "free/busy" (binary) status information.
- *Optical* IoT nodes (provided by things2i[5] and denoted as "park2i®") installed on top of the parking area to be monitored. This node provides data in two different modes, depending on the device configuration: (i) a list of "free/busy" (binary) statuses of parking spaces in the area; or (ii) the aggregated occupancy percentage of the entire parking area.

Beside the first PoC focused on parking sensors, the municipality of Parma started to deploy different IoT LoRaWAN devices, to cover different applications. In particular, the municipality identified the following five different "pillars" of interest.

---

[2] http://www.btenia.it/.

[3] This topic has been of particular interest for people visiting the city of Parma, Italy, which was the Italian Capital of Culture in the years 2020–2021: https://parma2020.it/en/.

[4] https://municipia.eng.it/prodotti/parking-spot-sensor/.

[5] https://www.things2i.com/.

Table 2 LoRaWAN-based IoT nodes currently deployed in the city of Parma, Italy, and integrated in the middleware

Type	Manufacturer	Model	Number of IoT nodes
Water metering	Midomet	Midomet LoRa Pulse/Analog	4
Parking monitoring	things2i	park2i®	4
Parking monitoring	Neosystems	Neosystems parking sensor (NPS)	39
Cars counter	Nablaquadro	LoRa traffic sensor	13
Bike and people counter	Parametric	Traffic sensor	3

- Large parking area monitoring (e.g., "park-and-ride" areas in the city, as well as supermarket parking lots) through the deployment of park2i® nodes.
- Single parking space monitoring, to monitor the use of specific or particular spaces (e.g., for handicapped people) through magnetic IoT nodes, as mentioned before.
- Mobility data collection through the deployment of bi-directional car, bike, and pedestrian counters along roads and bicycles lanes of particular interest for the municipality (e.g., limited-traffic zones and public bike paths).
- Water resource consumption monitoring in public buildings through the deployment of water metering sensors.
- Additional heterogeneous sensors, not related to a specific application but of future interest for larger deployments (e.g., noise sensors, trash volume and location sensors, etc.).

For the sake of clarity, in Table 2 we give a summary of all the LoRaWAN-based IoT nodes deployed in the city of Parma and integrated in the middleware.[6] The choice of the listed commercial IoT devices has been left to the municipality of Parma, which, together with its collaborators and providers, acquired and deployed the devices depending on the "pillars" of interest and the existing communication infrastructure.

Following the proposed smart city-oriented IoT middleware-based architecture, the *first integration step* required to add the IoT nodes deployed in Parma is the Node Registry module, storing all the details about each specific IoT device—such as manufacturer, model, serial number, transmission technology—and a reference to a specific device descriptor (in JSON format) containing the list of the resources monitored by the specific IoT device (e.g., temperature, humidity, passages, etc.), with a consequent mapping to one or more taxonomy entries known by the middleware itself. Then, the JSON descriptor is shared among all the IoT devices of the same type: as an example, with reference to the aforementioned parking monitoring scenario, two different descriptors should be added to the IoT architecture in order to describe the corresponding data sources. Another relevant piece of information stored in the Node Registry is the geographical location of the nodes (namely, GNSS-based latitude, longitude, and altitude values), required to link the sensed data with their respective location in Parma. Ultimately, this allows to build a "parking city overlay."

The *second integration step* is required to develop the connectors, in order to manage the different transmission modes and data rates. To this end, each connector has to acquire the data stream from its data source, based on the specific communication protocol. Then, exploiting (i) the JSON descriptors of each IoT device mapped in the architecture and (ii) the taxonomy definition, the Normalization module normalizes the data format, making it coherent for subsequent utilization in the middleware's processing units. As an example, with reference to the city of Parma, the data received by parking sensors are mapped using taxonomies such as `parking_status` (with `true` and `false`

---

[6] The IoT nodes listed in Table 2 refer to the time of writing, as their relative number and nature will be updated based on upcoming needs and future decisions.

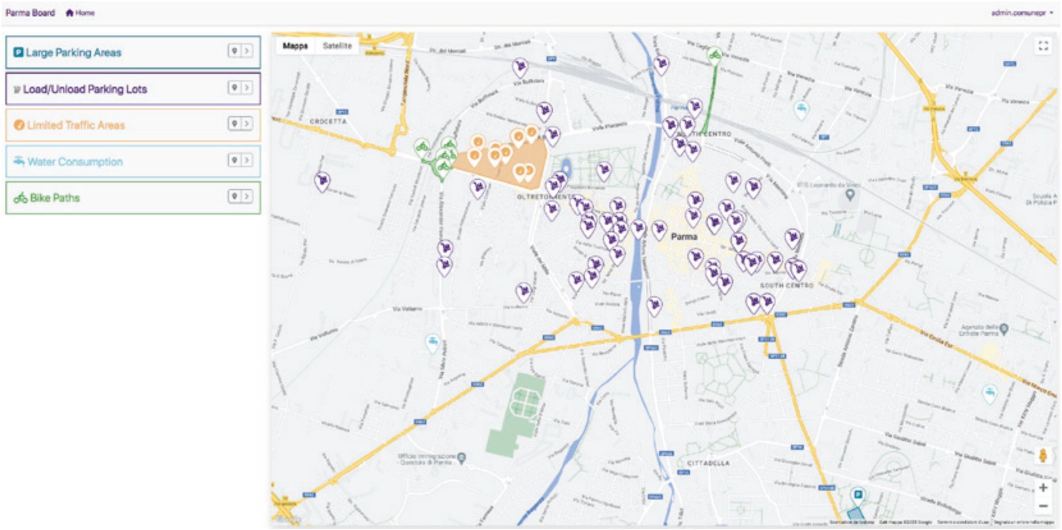


Fig. 2 Overview of the middleware data sources deployed in city of Parma, Italy

as possible values), whereas data arriving from pedestrian counters can be mapped with `left2right_counter`, `right2left_counter`, and `temperature` values (as these IoT devices have also an *on-board* temperature sensor).

At this moment, since in the city of Parma only LoRaWAN IoT devices (managed by a single LoRaWAN network operator) are considered in the PoCs, the following types of connectors have been developed.[7]

- HTTP Webhook connector, which *passively* receives data streams from the LoRaWAN Application Server (namely, through HTTP POST requests sent to the middleware's connector endpoint each time an IoT device sends a new LoRaWAN uplink message).
- MQTT connector, which *passively* receives data streams through a subscription on MQTT topics of interest, on which data will be published by the LoRaWAN Application Server.

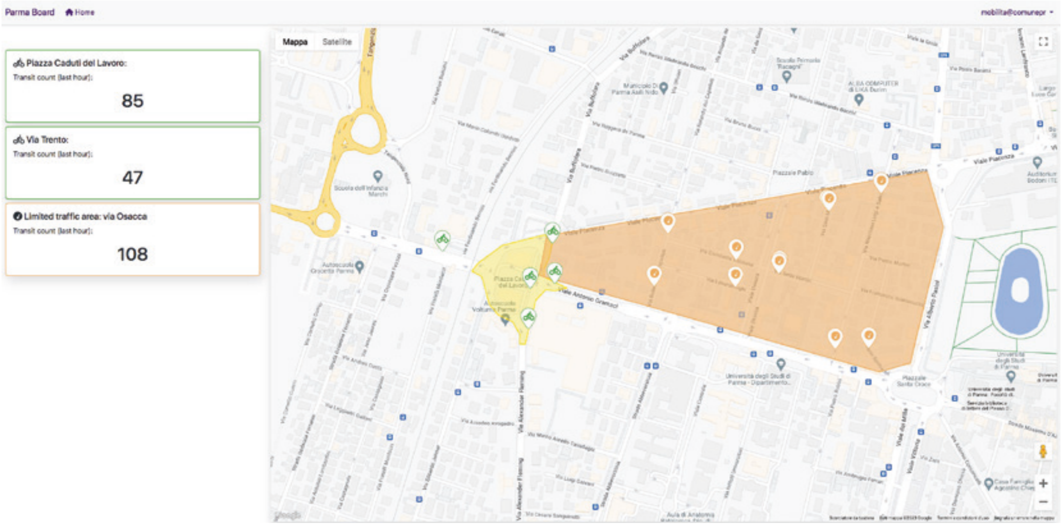
An illustrative representation of the proposed IoT-oriented middleware-based architecture applied to the city of Parma has already been shown in Fig. 1, where the deployed components and systems are depicted as colored clouds, forwarding their data through the pink solid lines.

Once the device integration in the IoT middleware has been completed, a comprehensive and unified view of the city status can be derived. In fact, it is possible to define a custom UI to show live data from a single IoT device or aggregated data from all available data streams with the same label (e.g., `parking_status`). No further considerations related to the specific deployment platform for the UI (e.g., desktop, mobile, Web, etc.) will be discussed here, as the *presentation layer* can be considered as an *add-on* for the consumers, as defined in the Dynamic Data Visualization module of the middleware. To this end, Fig. 2 shows a prototypical middleware Web-based dashboard for the city of Parma, providing an overview on the installed IoT devices and their location in the city, with reference to the "pillars" (or topics) of interest for the municipality.

The *last* step pertaining with the integration involves the Data Processing module, in turn using the integrated normalized data to perform complex operations. As an example, in the case of parking lots monitored with different IoT

---

[7] Additional connectors are under development based on the new data sources to be integrated.



Piazza Caduti del Lavoro:
Transit count (last hour):
85

Via Trento:
Transit count (last hour):
47

Limited traffic area: via Osacca
Transit count (last hour):
108

Fig. 3 Example of UI corresponding to complex overlays aggregated by specific middleware's processing units. In this case, the UI refers to a limited mobility area ("zona 30").

technologies, it is possible to create dynamic UIs based on the nature of the data sources, while developing custom data processing units for data fusion. In general, the middleware allows to activate processing units performing statistical operations on historical data, calculating values such as the average parking duration, the number of status changes and so on, as well as detecting correlations with other data overlays that can be available in the same city area. In the current deployment in the city of Parma the following types of processing units have been developed.[8]

- Counter data management: since the deployed car counter only provides an "absolute" count (starting from the time of installation), a processing unit has been developed to calculate the "relative" increments with respect to the time frame of interest.
- Limited-traffic-area monitoring: as shown in Fig. 3, in order to allow the municipality to benefit of a clear overview on the access to limited-traffic areas in the city, a processing unit (based on the knowledge of the locations of several car and bike counters) is in charge of computing the amount of vehicles currently present in specific areas.
- Data format conversions, or geographical filtering: in these cases, the developed processing units are in charge of uniforming the notation of the information received from data producers, or returning a data subset based on a specific geographical area selection, respectively.

The first implementation of the middleware has been deployed in 2022 and has been collecting data for about 8 months, with increasing number of data sources. Being a modular architecture, its performance is difficult to evaluate, since it depends on several aspects. Despite the complexity, some preliminary considerations can be made. The data acquisition time, defined as the time elapsed from the generation time (declared by external data sources, e.g., the LoRaWAN Application Server) to the time when the information is stored in the middleware and associated with a taxonomy entry, is shorter than 1 s in over 98% of the acquisitions, with an average input data rate of 13 new data entries per minute. This acquisition time, however, can increase

---

[8] As for the additional connectors mentioned before, newly specific processing units are under development based on the needs of the municipality, in terms of enhanced services to be provided to the end-users.

in an unpredictable way if data sources require custom processing in the *Data Processing module*.

Moreover, the proposed middleware-oriented architecture is designed to integrate streams of heterogeneous data sources (from both IoT devices and external software tools). In particular, the maximum number of sources that the system can handle depends on several factors, such as:

- the data generation rate (e.g., LoRaWAN devices should abide by duty cycle constraints);
- the payload size;
- the paradigm or custom processing required for data acquisition (e.g., connectors for external software can follow a polling paradigm through REST API, but also a publish/subscribe model).

To this regard, the proposed architecture is highly modular, making it possible to add new components depending on the workload (e.g., add a specific virtual machine to handle the data flows of a specific connector). The cost is directly related to the complexity of the final architecture and can be billed to the municipality as a periodic subscription fee.

Even though the IoT deployment in the city of Parma is still at an early stage, the proposed middleware-oriented architecture represents a useful tool with wide applicability, since it hides all the technical details related with the deployment of heterogeneous devices (each one with its own technology), which may not be of interest for the municipality. Moreover, it allows to build an urban dataset without the need to access vertical applications by different manufacturers. The collected information can then be forwarded to external entities through specific data connectors. For example, this may include historical data in a specific time range, made accessible through HTTPS REST APIs for statistical purposes, as well as *live* data forwarded through a WS-based channel to developers working on mobile Apps for citizens.

4 Conclusions and Future Integration

The diffusion of IoT technologies in urban contexts, together with data analysis (also based on AI techniques), plays a key role in the implementation of the concept of smart cities aiming at enhancing the citizen's quality of life and improving the city's administrative processes. IoT often pertains to different domains, such as: environment and pollution monitoring, health care and education, energy and resources management, and mobility. Unfortunately, the lack of shared and common standards and platforms has generated the diffusion of proprietary and disjoint vertical applications. These verticals allow to collect large amounts of data from a city but, at the same time, are not designed to communicate with each other for data fusion. This chapter discussed a smart city IoT-oriented middleware-based architecture, whose aim is integrating heterogeneous data generated from multiple sources, in order to allow a municipality to monitor the city status from multiple perspectives. This allows to highlight (sometimes discover) "hidden" correlations among heterogeneous IoT data by building synergies between isolated systems. The core of the middleware is the utilization of a shared taxonomy and the concept of connectors, which perform the following operations: data collection, data format translation, and data normalization for each class of data sources (IoT systems as well as external software tools). The modularity and scalability of the proposed IoT architecture allows to hide the heterogeneity of possible data sources and technical aspects of the adopted communication protocols, thus providing a city control panel for the municipality to monitor processes and to build new services that can simplify the city management, and optimize the citizens' and visitors' experience. Finally, for the sake of validation of the proposed IoT middleware-oriented platform, its actual experimental deployment in the city of Parma, Italy, has been discussed.

References

1. D. Belleri, M. Baick, C. Ratti, J. e-Learn. Knowl. Soc. **18**(3) (2022). https://doi.org/10.20368/1971-8829/1135820
2. S.E. Bibri, J. Krogstie, Sustain Cities Soc. **31**, 183 (2017). https://doi.org/10.1016/j.scs.2017.02.016
3. P. Bellini, P. Nesi, G. Pantaleo, Appl. Sci. **12**(3), 1607 (2022)
4. L. Belli, A. Cilfone, L. Davoli, G. Ferrari, P. Adorni, F. Di Nocera, A. Dall'Olio, C. Pellegrini, M. Mordacci, E. Bertolotti, Smart Cities **3**, 1039 (2020). https://doi.org/10.3390/smartcities3030052
5. T. Abbiasov, C. Heine, E.L. Glaeser, C. Ratti, S. Sabouri, A. Salazar Miranda, P. Santi, The 15-Minute City Quantified Using Mobility Data. Working Paper 30752, National Bureau of Economic Research (2022). https://doi.org/10.3386/w30752
6. L. Sanchez, L. Muñoz, J.A. Galache, P. Sotres, J.R. Santana, V. Gutierrez, R. Ramdhany, A. Gluhak, S. Krco, E. Theodoridis et al., Comput. Netw. **61**, 217 (2014)
7. S. Latre, P. Leroux, T. Coenen, B. Braem, P. Ballon, P. Demeester, in *2016 IEEE International Smart Cities Conference (ISC2)* (IEEE, 2016), pp. 1–8
8. X. Liu, M. Chen, C. Claramunt, M. Batty, M.P. Kwan, A.M. Senousi, T. Cheng, J. Strobl, A. Coltekin, J. Wilson, T. Bandrova, M. Konecny, P.M. Torrens, F. Zhang, L. He, J. Wang, C. Ratti, O. Kolditz, A. Klippel, S. Li, H. Lin, G. Lu, The Innov. **3**(5), 100279 (2022). https://doi.org/10.1016/j.xinn.2022.100279
9. R. Kashyap, M. Azman, J.G. Panicker, in *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)* (IEEE, 2019), pp. 1–5
10. L. Davoli, G. Ferrari (eds.), *Wireless Mesh Networks for IoT and Smart Cities: Technologies and Applications*. Telecommunications (Institution of Engineering and Technology, 2022). https://doi.org/10.1049/PBTE101E
11. A. Cilfone, L. Davoli, L. Belli, G. Ferrari, Future Internet **11**(4), 99 (2019). https://doi.org/10.3390/fi11040099
12. D. Magrin, M. Centenaro, L. Vangelista, in *2017 IEEE International Conference on Communications (ICC)* (2017), pp. 1–7. https://doi.org/10.1109/ICC.2017.7996384
13. L. Feltrin, G. Tsoukaneri, M. Condoluci, C. Buratti, T. Mahmoodi, M. Dohler, R. Verdone, IEEE Wireless Commun. **26**(1), 78 (2019). https://doi.org/10.1109/MWC.2019.1800020
14. L. Davoli, L. Belli, A. Cilfone, G. Ferrari, IEEE Internet Things J. **5**(2), 784 (2018). https://doi.org/10.1109/JIOT.2017.2747900
15. M. Cesana, A.E.C. Redondi, *IoT Communication Technologies for Smart Cities* (Springer International Publishing, Cham, 2017), pp. 139–162. https://doi.org/10.1007/978-3-319-44924-1_8
16. I. Yaqoob, I.A.T. Hashem, Y. Mehmood, A. Gani, S. Mokhtar, S. Guizani, I.E.E.E. Commun, Magaz. **55**(1), 112 (2017). https://doi.org/10.1109/MCOM.2017.1600232CM
17. A. Haidine, S. El Hassani, A. Aqqal, A. El Hannani, in *Smart Cities Technologies*, ed. by I. Nunes Da Silva, R.A. Flauzino (IntechOpen, Rijeka, 2016), Chap. 4. https://doi.org/10.5772/64732
18. NB-IoT vs LTE-M: A comparison of the two IoT technologies. https://onomondo.com/resource-hub/nb-iot-vs-lte-m-a-comparison-of-the-two-iot-technology-standards/. Accessed on 17 Mar 2023
19. B.S. Chaudhari, M. Zennaro, S. Borkar, Future Internet **12**(3) (2020). https://doi.org/10.3390/fi12030046
20. J. Gozalvez, IEEE Vehicular Technol. Magaz. **11**(1), 14 (2016). https://doi.org/10.1109/MVT.2015.2512358
21. L. Davoli, I. Paraskevopoulos, C. Campanella, S. Bauro, T. Vio, A. Abrardo, G. Ferrari, Sensors **21**(4), 1329 (2021). https://doi.org/10.3390/s21041329
22. S.K. Rao, R. Prasad, Wireless Personal Commun. **100**(1), 161 (2018). https://doi.org/10.1007/s11277-018-5618-4
23. I. Yaqoob, I.A.T. Hashem, Y. Mehmood, A. Gani, S. Mokhtar, S. Guizani, I.E.E.E. Commun, Magaz. **55**(1), 112 (2017). https://doi.org/10.1109/MCOM.2017.1600232CM
24. M. Bauer, L. Sanchez, J. Song, Sensors **21**(13), 4511 (2021)
25. J. Koo, Y.G. Kim, in *Proceedings of the 36th Annual ACM Symposium on Applied Computing* (2021), pp. 690–698
26. R.J. Hassan, S.R. Zeebaree, S.Y. Ameen, S.F. Kak, M.A. Sadeeq, Z.S. Ageed, A. AL-Zebari, A.A. Salih, Asian J. Res. Comput. Sci. **8**(3), 32 (2021)
27. J.J. Peralta Abadía, C. Walther, A. Osman, K. Smarsly, Sustainable Cities and Soc. **83**, 103949 (2022). https://doi.org/10.1016/j.scs.2022.103949. https://www.sciencedirect.com/science/article/pii/S2210670722002700
28. Q. Alfalouji, T. Schranz, A. Kümpel, M. Schraven, T. Storek, S. Gross, A. Monti, D. Müller, G. Schweiger, Buildings **12**(5), 526 (2022)
29. J.N.S. Rubí, P.R. de Lira Gondim, International Journal of Communication Systems **34**(2), e4515 (2021)
30. J. Pereira, T. Batista, E. Cavalcante, A. Souza, F. Lopes, N. Cacho, Future Generation. Comput. Syst. **128**, 552 (2022) https://doi.org/10.1016/j.future.2021.10.030. www.sciencedirect.com/science/article/pii/S0167739X2100426X
31. Z. Shelby, K. Hartke, C. Bormann. Rfc 7252: The constrained application protocol (coap) (2014)

32. L. Davoli, L. Belli, G. Ferrari, Enhancing Security in a Big Stream Cloud Architecture for the Internet of Things Through Blockchain (IGI Global, Hershey, PA. Chap. **5**, 104–133 (2019). https://doi.org/10.4018/978-1-5225-8295-3.ch005
33. L. Belli, S. Cirani, L. Davoli, G. Ferrari, L. Melegari, M. Picone, Applying Security to a Big Stream Cloud Architecture for the Internet of Things (IGI Global. Chap. **57**, 1260–1284 (2020). https://doi.org/10.4018/978-1-5225-9866-4.ch057
34. L. Belli, S. Cirani, A. Gorrieri, M. Picone, in *Proceedings of the 1st International Workshop on Experiences with the Design and Implementation of Smart Objects* (Association for Computing Machinery, New York, NY, USA, 2015), SmartObjects '15, pp. 1–6. https://doi.org/10.1145/2797044.2797046