An IoT Framework for Heterogeneous Multi-layered Access in Smart Cities

Evangelos Syrmos[1](✉), Dimitrios Bechtsis[1], Ioannis Tsampoulatidis[2,3], and Nicos Komninos[3]

[1] Department of Industrial Engineering and Management, School of Engineering, International Hellenic University, Thessaloniki, Greece
syrmevag@iem.ihu.gr
[2] Infalia PC, Thessaloniki, Greece
[3] URENIO Research, Faculty of Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece

Abstract. The proliferation of the Internet of Things has been instrumental in the digitalization of smart cities, where various technologies are leveraged to enable data utilization. However, achieving interoperability among diverse technologies remains a challenge due to heterogeneity. In this regard, ontologies have been proposed as a standalone solution that semantically enriches IoT data. Nevertheless, ontologies are still underutilized due to several limitations as systems become more complex. In order to overcome these limitations and provide a single point-of-access for all smart city layers, this paper presents a theoretical IoT framework that minimizes the requirement of interdisciplinary knowledge to operate IoT platforms, while utilizing existing ontologies in a complementary manner to build a high-level ontology schema. The framework is comprised of five fundamental axes/pylons that enable frictionless usage and configuration in order to extract data from all employed smart city layers via an adapter. This can be beneficial for smart cities such that legacy or under-utilized IoT platforms can be integrated with the framework and provide additional information reducing the costs. External operators, applications and platforms can configure the axes in order to extract data regarding their needs and make better decision making.

Keywords: IoT · Framework · Smart Cities · Ontologies · Interoperability

1 Introduction

IoT is becoming the core technology for enabling message broadcasting and inner communication between all smart city ecosystems. Therefore, smart cities (SC) consist of a plethora of sub-ecosystems (domains) that handle specific needs, however, integrating diverse technologies in order to achieve interoperability is challenging [1]. Initiatives that foster interoperability and decrease heterogeneity have been at the forefront of the research community [2]. Interoperability is a challenging facet across diverse SC ecosystems that utilize different technologies to generate, transform, broadcast and store data [3].

N. A. Streitz and S. Konomi (Eds.): HCII 2023, LNCS 14037, pp. 156–171, 2023.
https://doi.org/10.1007/978-3-031-34609-5_11

The maturity of ontologies enabled the categorisation of objects and and their surrounding environments in hierarchical trees by highlighting specific descriptions and functionalities. Ontologies have been proposed in frameworks that strive for interoperability [4]. From the IoT perspective, however, ontologies have been utilized as a classification and semantic enrichment tool on data streams [5]. Ontologies such as Sensors, Observation, Sample, and Actuator (SOSA), and Semantic Sensor Network (SSN) are used for providing additional meta-information regarding IoT sensors [6], however, Gonzalez-Gil state that although ontologies are beneficial, limitations must be considered thoroughly [7]. When scaling is required, the complexity of integrating ontologies significantly increases, due to the lack of automated relationship mechanisms, while the lack of a single framework that tackles the fragmentation and heterogeneity leads to a high maintenance cost of IoT systems. Interoperability between different ontologies is also a requirement even in a single domain as different definitions for the same characteristics may exist. Moreover, semantic heterogeneity could hinder the adoption since the representation of ontologies can be different while the usage from IoT stakeholders can lead to semantic ambiguity [8, 9]. Performance is also considered a limitation in ontology adoption, especially in real-time and resource-constraint devices. In order to strengthen the operational capacity of SC stakeholders, heterogeneous IoT devices with diverse specifications and features should have as a basis a common understanding of the context of their shared data. To this end, the proposed IoT framework addresses heterogeneity in the IoT ecosystem by introducing 5 pillars as common scales of measurement while proposing an ontology that leverages two axes for a double alignment approach. The framework proposes a junction of two solutions in order to reduce heterogeneity and improve interoperability. A four-stage ontology integration and a central-hub acting as a domain-neutral ontology are proposed for utilizing existing domain-specific ontologies whilst introducing a hierarchical observation for seamless discovery.

Proprietary IoT platforms and/or applications that interface with the IoT devices have been developed to provide easy access to city officials. However, integrating IoT platforms from various domains and SC ecosystems in a single point-of-access is challenging due to the demand of high technical literacy [10]. The proposed IoT framework presents the fundamental axes that enable operators to interact and retrieve information across all SC ecosystems. This abstraction from domain-specific IoT platforms reduces the necessity for city operators to possess direct access and familiarity with each constituent system in each SC sub-ecosystem. In order to facilitate the extraction of relevant information abstraction can ultimately improve decision-making. Existing platforms can integrate with the framework via the adapter layer while legacy applications can be retrofitted. This approach reduces the technical debt from underutilized platforms due to a lack of specialized personnel. The point of the matter is that the proposed framework minimized the requirement of interdisciplinary knowledge to operate IoT platforms since the adaptation of data to the presented framework is handled by domain experts.

The remainder of the paper is structured as follows: The fundamental axes/pylons of the theoretical IoT framework are presented in Sect. 2. Section 3 highlights the impact of the framework on SC layers. In Sect. 4, a detailed analysis of the data alignment from diverse urban datasets and ontologies is presented. Section 5 provides a use case scenario that leverages the theoretical framework. Finally, Sect. 6 concludes the paper.

2 Fundamental Axes

It is evident that IoT device integration and deployment in SCs is increasing at a rapid pace, which imposes several challenges. Smart sensors and actuators that enable real-time sensing capabilities significantly differ inside a specific SC layer. Therefore, device interfacing for data extraction demands specialized knowledge and in-depth domain understanding. For instance, a smart water metering solution integrated into the infrastructure layer that measures real-time water consumption demands different interfaces compared to a smart sensor that monitors traffic congestion on a highway [11–13]. Ontologies can provide the necessary object descriptors for capturing functionality traits and features. However, ontologies lack scalability when a high-level observation of multi-layered data aggregation is required, due to the fact that every ontology has different descriptors for a single object. Post-correlation and mapping of aggregated data requires an interdisciplinary approach to develop a high-level ontology schema. Moreover, this approach hinders scalability when new objects and smart city layers are included.

To this end, a theoretical framework is presented, addressing the scalability concerns by having five fundamental axes as context descriptors for the IoT devices. This design facilitates seamless interaction with all SC layers by using each axis as a configurable option that enables frictionless data aggregation across the smart city ecosystem. The framework presents the following axes:

1. Temporal
2. Spatial
3. Variation
4. Intensity
5. Edge capabilities

As shown in Fig. 1 the framework's architecture addresses scalability and data handling activities in the SC layers by proposing a five-axes framework, a meta-ontology along with an adapter layer. The framework itself is responsible for receiving incoming requests from external applications and modifying the requests based on the five available axes. Thereafter, according to the five provided axes, a final query is created that is passed through a meta-ontology and the adapter layer before accessing each SC layer separately.

The auditability provided by the framework is considered a key enabler in interfacing with the system across multiple smart city layers without requiring background knowledge of the underlying applications. However, for enhanced applicability and reliability, the meta-ontology adapter is responsible for interfacing with every layer. Providing specific axes as interface configurators facilitates the abstract representation of the interconnected sub-systems of each SC ecosystem.

External applications request data in order to identify useful information patterns. The adapter is responsible for interfacing with individual systems in every SC layer. Existing systems developed by domain experts are required to expose connectivity solutions that interface with the adapter. Which expedites the development processes and removes interdisciplinary overhead. The meta-ontology aggregates information from multiple ontologies that are included at every SC layer and propagates datastreams to

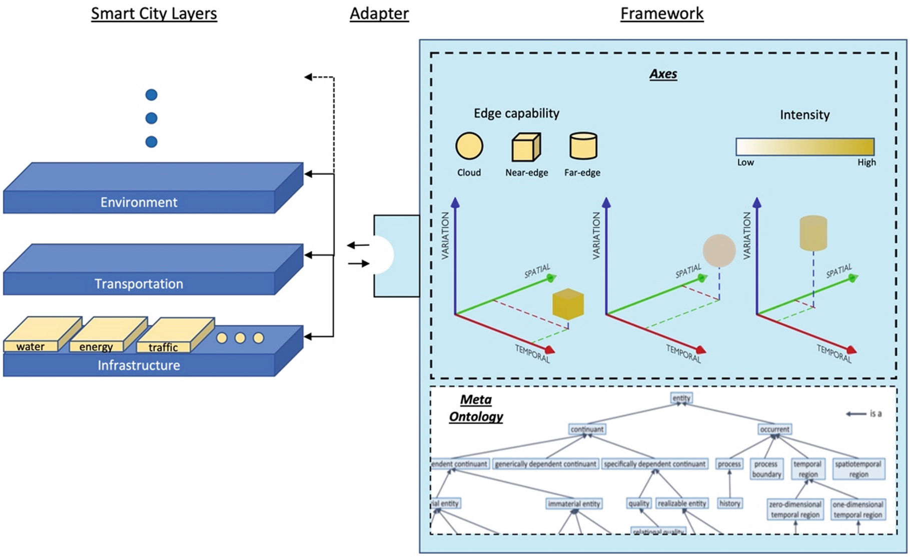

Fig. 1. High-level architecture and fundamental axes/pylon of the framework

the framework. Finally, the five axes are able to finetune multiple data streams using the axis boundaries.

A variety of configurations can be effectively established by the incoming queries. For instance, the request can be translated into a variety of configurations that adjust accordingly to every axis as presented in Fig. 1.

Having five axes to configure provides all the necessary means of extracting data across all involved smart city layers while combining information across multiple axes can enable further feature extraction and shed light on unknown areas. A theoretical description of each axis/pylon is presented along with a query example for better comprehension.

2.1 Temporal

The temporal axis, has a crucial role in the proposed theoretical framework. Time is a fundamental aspect during the generation and processing of information by IoT devices as data streams are constantly being generated. Specifically, the ability to capture and annotate data on the fly is essential for effectively understanding and utilizing them for further operations.

To facilitate this, datastreams are timestamped by the devices themselves which enable the comprehension of the fluctuation of data. Data sampling rates are also critical for estimating specific characteristics. The correlation of the timestamps, the sampling rates and the measured values with the data streams contributes to identifying patterns, trends and anomalies that would not be apparent without considering the temporal aspect of the data.

Furthermore, this axis also enables the development of time-sensitive decision-making, such that the system is able to react to the changing environment in real-time.

IoT relies upon timestamped data since it is a vital aspect of the systems to adapt and respond.

From the operator's perspective, having the temporal axis as a configurable option enables data aggregation across all smart city layers that match the given request. For instance, the request can be "fetch all data across all smart city layers that were created in the past day having either 1-min or 1-h intervals".

2.2 Spatial

Similarly, the spatial axis is also considered a fundamental aspect in the IoT domain. This axis pertains to the physical location of the employed IoT devices and is crucial in understanding the relationship between the devices and their sensing location.

Operating IoT devices that capture data in the field by sensing changes in the environment are typically assigned a dedicated geographic location which enables the correlation between the generated data, the exact location of the device or the wider area of interest. The significance of the spatial aspect is particularly pronounced in IoT solutions where device location does not change. In order to achieve locational homogeneity across different data layers, SC can use a geometrical grid as a layout for data recording reference where grid cells can act as spatial entities. This common spatial reference approach can ultimately reduce spatial fragmentation.

In these cases, IoT management platforms running on the cloud are responsible for registering device information including the. Furthermore, these platforms annotate incoming information based on the broadcasted device ID and their registered location before data are being stored in the database. On the other hand, devices that change their location constantly are equipped with specialized hardware (such as GPS antennas) that track the geographic location and annotate data before transmission.

Operators can request information based on certain criteria e.g., fetch all data across all smart city layers that are located in a specific area between latitude-A, longitude-A and latitude-B, longitude-B, this query will return all devices that are within the specified area, while the framework could aggregate the measurements and compute the mean value for this area.

2.3 Variation-Fluctuation

Having the ability to configure the aggregated data based on sensor variations is critical for enabling domain experts to evaluate each aspect. This axis presents the ability to map data fluctuations across all involved smart city layers. Hence, bridging the information about the dynamic nature of IoT systems that monitor the environment. For instance, a deployed temperature sensor monitors the ever-changing value throughout the day. Values such as temperature, and humidity have low fluctuations while a traffic sensor that monitors the pedestrian or vehicle flows can be categorized as a high variation device. Therefore, this axis can be considered complementary since high-fluctuating devices can leverage the ability of the system to detect critical events or anomalies.

2.4 Intensity

Intensity is another aspect of configuring the query based on the amount of generated data across all smart city layers. Specifically, intensity is directly related to the temporal axis/pylon, however, the frequency of data generation is promoted by employing intensity as an axis/pylon. For instance, operators might request data with a frequency of 5ms without hard-lined boundaries on time, thus forcing the theoretical framework to request data that have the ability to transmit and sample data at a such high rate. Similarly, combining intensity with spatial axis can ultimately provide direct insight into generated data in a given geographic location. The requested location can also be identified as a high or low-intensity area that is being monitored. Smart cities can benefit from this axis by identifying the IoT readiness of the smart city. Intensity and IoT-readiness are intertwined due to the fact that high intensity can be achieved purely on high digitalization and IoT adoption. Enforcing digitalization across all districts in smart cities will indisputably lead to a highly digitalized city that operates with minimal human intervention. Smart cities fed with high frequency and variety of data across all layers facilitate proactive behaviours, while underdeveloped areas can be strengthened to enable digital transformation.

2.5 Edge Capability

With the rising adoption of edge computing in the IoT landscape, we opted to include edge computing in the proposed theoretical framework.

Far-edge is the least compute capable category by all means, due to the fact that constrained devices are included. These devices lack computing resources and processing power required to perform intensive operations. In most cases, these devices are used for integrating multiple sensors that require external computing units to monitor the environment. For instance, LoRaWAN end nodes which are required to operate on batteries for multiple years are not able to perform on-site computation on the aggregated data.

On the other hand, the near-edge is considered a more capable category for performing highly optimized algorithms and machine learning techniques. Gateways and hubs consolidate the edge layer since these devices are required to accommodate multiple transmitting devices on multiple channels at the same time. Of importance to understand in this regard is that these devices are able to process incoming data by multiple devices in real-time, and perform data-cleaning, formatting and filtering with intelligent algorithms or even with machine learning models. However, deploying machine learning models on such devices requires decreasing the overall size of the model while sacrificing the accuracy in order to be utilized by the device.

The last category is the cloud, which is mostly directed to the platforms and/or applications that act as data aggregators. These have the ability to perform heavy workloads and intensive operations on all data without affecting accuracy and performance. Most of the time these are located on servers with highly capable hardware.

Concluding, this axis/pylon has been introduced mainly for providing operators with the flexibility to query information about the employed devices accordingly. For instance, operators can request for all information across all smart city layers that have far-edge computing capabilities. This will return all device values that meet the specified criteria.

3 Axes' Impact on Smart City Layers

The proposed framework presents several key benefits to SCs by improving the overall efficiency and effectiveness of process management activities. The five fundamental axes foster the datastream alignment that decreases the fragmentation and heterogeneity aspect. Moreover, it enables the use of datastreams based on specific scales of measurement in a homogenized format. Specifically, the framework provides a single point of access to multiple platforms that have been developed and underutilized. This streamlines the data extraction process which ultimately reduces the friction between data aggregation from different sources across all SC layers. This not only reduces the personnel requirements but also enables non-specialized operators to access data from multiple city ecosystems, align them on a single context vertical that supports the decision-making process. Urban planners can make forecasts and better decisions on resource management, while real-time event handling can greatly be improved.

Another highly influential aspect of the framework is that data extraction across all city layers provides urban planners with a holistic view of the city, allowing them to make informed decisions and implement effective strategies. Moreover, data correlations, patterns and forecasts can be observed across different city ecosystems which by design is difficult to achieve due to high complexity.

The framework also contains a top-level ontology that is embedded in a central hub in order to leverage two of the fundamental axes as alignment planes. The use of a top-level hierarchical domain-neutral ontology can be easily adopted by completely different cities. Due to the fact that a spatial and temporal alignment acts as the driving verticals that facilitate the representation of data across multiple data layers on a common reference entity.

The provided flexibility to configure the data extraction process according to specific requirements is possible via the given axes. It is crucial to ensure that the framework is tailored to meet the needs of each city, maximizing its overall effectiveness. Consequently, urban planners can select the data streams and the accuracy levels within the framework (positioning the data stream within the five axes) to improve the overall effectiveness.

Another key impact of the framework is that it enables the connectivity, and integration of individual applications, reducing the overall complexity and fragmentation of existing platforms. Underutilized applications can also be retrofitted into the framework which can increase the Return on Investment (ROI) and increase frameworks' reach on uncharted areas.

4 Data Alignment Across Urban Datasets and Ontologies

4.1 Pooling of Urban Datasets: Fragmented and Siloed Information

Following the multi-level observation and data collection we described in the previous sections, a pool of datasets that depict a city or its ecosystems is created. This data pool is fragmented and heterogeneous. Heterogeneity is due to different sources of data, different organization principles, scales of measurement, and many other features specific to each data source used. Consequently, it becomes particularly difficult to get a comprehensive view of the available data assets, reconcile them, and get meaning about the realities of the city these data capture and represent. There are two types of fragmentation and heterogeneity in a dataset created from the juxtaposition of multiple data sources, layers of observation, and recording systems.

First, there is a spatial or temporal fragmentation of the data as they are recorded along different spatial entities or different time periods. For instance, building-related data can be recorded using the "building" as a spatial entity of reference; land use data can be recorded using the "building block" as a spatial entity of reference, and infrastructures can be recorded using the "city grid" as a spatial entity of reference. The dimensions and location of these spatial entities differ substantially (Fig. 2). However, we need to achieve locational homogeneity of data by using common spatial entities of reference across different data layers. We may use the city grid and aggregate buildings, land uses, activities, and other data at this spatial level. Or we may plot a geometrical grid over a city and use the grid cells as spatial entities of reference for data recording. No need to say that locational homogeneity is important if we wish to represent the realities of a city in a comprehensive way. The same fragmentation applies to temporal data if the recording of different categories of data took place in different time periods.

Fig. 2. Spatial units at the level of the building (red), building block (green), and city grid (blue)

Second, there is semantic heterogeneity as data capture and represent different aspects of the urban reality along different taxonomies. Whatever the spatial entity of reference, data recording concerns data related to (a) the landscape and physical environment of cities, such as buildings, constructions, infrastructures, and natural ecosystems, (b) the social characteristics of cities, such as land uses, activities, events, communities and groups, crime, etc. (c) the digital infrastructure of cities, broadband networks, sensors, IoT, websites and social media, and (d) the population, age groups, employment, income and wealth, and many other characteristics [14]. Each category of data is described by a taxonomy or ontology, which is a description of data structure with classes, properties, relationships, and axioms. Taxonomies offer the basis to ensure both data consistency and understanding of the underlying data model. For instance,

- buildings can be described with the Building Topology Ontology, a minimal OWL DL ontology, having as classes those of building, element, interface, site, space, storey, and zone [15]
- conditions of the environment can be described with the Web of Things (WoT) ontology [16], or the Thing Description (TD) Ontology [17]
- land uses can be described with the Land Use Ontology that captures types of land use and cover over time [18],
- activities can be described with the NACE Rev. 2 standard classification of economic activities [19],
- risks and city threats can be described by using a risk ontology [20].

Semantic alignment and content correlation are important because they give meaning to data. A bus station in the freezing north is not the same as a bus station in the Mediterranean; a school in a high-crime area is not the same as a school in a zero-crime village. Present knowledge changes past knowledge, offering new interpretations of past events. Correlations give new meaning, which is lost when data is fragmented and heterogeneous. However, the integration of content in datasets created along different ontologies is more challenging than their spatial integration.

4.2 Methods to Align Heterogeneous Datasets

Integrating data from heterogeneous sources and making queries is an important topic in database design, multi-disciplinary engineering, semantic web applications, and elsewhere, having as an objective to provide comprehensive access to heterogeneous data sources. Some methods use ontologies and RDF schemas to represent content from heterogeneous sources.

Osman, Yahia, and Diallo address the heterogeneity problem through the integration of existing ontologies to build a new more coherent one [21]. They offer an overview of the literature with the most relevant works in the field, key definitions of ontology integration, integration principles, consequences, and techniques. Ontology integration can be split into repairing, matching, and merging steps, each one preparing the terrain for the next step. The key process is ontology matching, also referred to as ontology alignment, which consists in establishing semantic correspondences between the entities of different ontologies, mainly matching classes and properties of different ontologies. Ontology alignment is the outcome of this matching process. Following the authors, the ontology integration workflow is developed in four phases: (a) pre-processing, to analyse input ontologies and improve their quality in order to reduce the matching effort, (b) matching, which identifies correspondences between the input ontologies and pairs of equivalent entities across the ontologies under consideration, (c) merging with inputs from all entities into a new integrated ontology, and (d) post-processing, to assess, repair and refine the resulting new ontology. Throughout this paper, they provide exhaustive matching, alignment, and correspondence types, as well as types of ontology merging and ontology integration.

A similar review can be found in the paper of Ekaputra et al. on ontology-based data integration in multi-disciplinary engineering environments [22]. They report on 23 applications from both the semantic web and system engineering and identify authors and methods, integration variants, key problems, strengths, and limitations of the different integration approaches.

Wang et al. present a different approach based on a mediator-wrapper architecture that includes four layers: (1) an application layer to communicate with users; (2) a mediating layer to perform data integration; (3) a wrapper layer which contains wrappers for data sources; and (4) a source layer giving access to heterogeneous data sources [23]. The core component in the proposed solutions to solve the heterogeneity problem is a relational schema based on an entity-relationship diagram (ERD), which interconnects tables of a relational database over an ontology and RDF associations (Fig. 3).

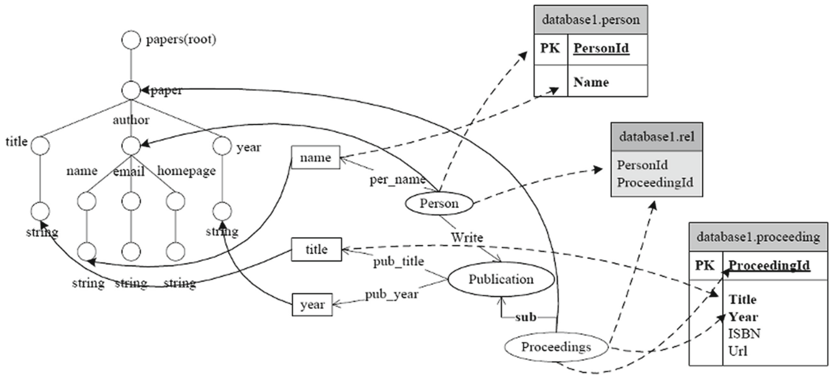

Fig. 3. Semantic mapping between relational schemas and ontology [21]

An earlier paper by Dou and LePendu, discussed ontology-based integration for relational databases [24]. For them "a merged ontology is the ontology equivalent of a global view over local schemas. It consists of common elements from a source ontology and a target ontology but also defines the semantic mappings between them as bridging axioms. A merged ontology allows all the relevant symbols in a domain to interact so that facts can be translated from one ontology to another using inference over the bridging axioms". They argue that defining semantic relationships between concepts is too subtle for full automation and human interaction is needed. Thus, they define the bridging axioms manually, based on the understanding of semantic relationships between the local schemas.

Elmhadhbi, Karray, and Archimède show a different path in which semantic interoperability across different systems and information sources can be achieved by aligning to upper-level ontologies to come up with a shared vocabulary and understanding [25]. They present a use case in which the Basic Formal Ontology (BFO), an upper-level ontology, and the Common Core Ontology (CCO), a mid-level ontology, are combined to define a new ontology for firefighters composed of 429 classes and 246 relations. The upper-level ontology helped to improve data quality, to reduce development time and especially facilitate information integration, avoid inconsistencies and achieve both syntactic interoperability to exchange information and semantic interoperability to ensure that information exchanges make sense under a common understanding.

The above literature briefly presented shows that merging ontologies is a largely manual and time-consuming process. It contains aspects that are equivalent to building a new ontology from scratch, using classes, instances, object properties, annotations, data properties, and axioms of the heterogeneous datasets under consideration. Eventually, depending on the number of ontology entities to be integrated, building a new ontology might be less time-consuming and more coherent.

4.3 Integration and Alignment of Fragmented Urban Datasets

The path we propose to integrate heterogeneous datasets for a city organized by a group of different ontologies combines two approaches:

- The four-stage workflow for ontology integration outlined by Osman, Yahia, and Diallo enables to set of semantic correspondences and matching between entities of different ontologies [19],
- The use of a top-level ontology as a central hub to align spatial and temporal classes based on the presented axes and create new relationships from the initial ontologies [26].

The process is depicted in Fig. 4 in which a group of heterogeneous datasets and their underlying ontologies (DS-O1, DS-O2, DS-O3, DS-O4) evolves into a hub and spokes architecture to create a suite of interoperable ontologies. The process includes the setting of a top-level ontology at the central hub and the alignment/transformation of initial datasets (DS-O1$_{TR}$, DS-O2 $_{TR}$, DS-O3 $_{TR}$, DS-O4 $_{TR}$).

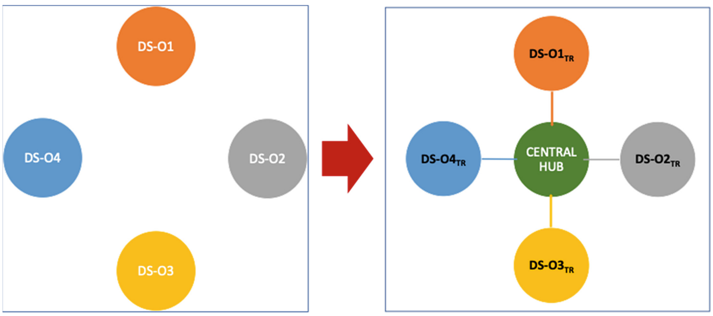

Fig. 4. Integration and alignment of heterogeneous datasets and ontologies

Three objectives guide this double alignment process: (a) the need to create a new hierarchy over all data classes included in the initial datasets, (b) the need to align the spatial and temporal classes across different ontologies, and (c) the need to introduce new properties connecting instances across the ontologies used, thus creating new meaning at the intersection of initial ontologies.

The Central Hub. A domain-neutral ontology is developed in the central hub to connect with domain-specific ontologies and support retrieval and discovery throughout the datasets. ISO/IEC 21838–1:2021 specifies the requirements for a top-level hub ontology, defining the relations between top-level ontology and domain ontologies and the role of a top-level ontology in definitions and axioms of domain-specific ontologies [27]. As a top-level ontology at the central hub, we propose the BFO-ISO, an ontology evolved from the Basic Formal Ontology (BFO) and defined as top-level ontology by ISO 21838–2 [28]. BFO is widely used to facilitate interoperability across multiple engineering-related ontologies; is a realist formal ontology representing high-level universal types of things; does not contain any domain-specific knowledge [29, 30]. We have used BFO to develop the SC ontology and provide a better understanding and description of the smart/intelligent city landscape, identify main components and processes, and clarify core entities related to the integration of physical, social, and digital dimensions of a city [31].

The class hierarchy of BFO-ISO: ISO 21838–2 includes a limited number of entities. It starts with the dichotomy between *continuants* (material or immaterial entities that continue to exist through time while maintaining their identity) and *occurrents* (procedures that unfold over a time period); adopts the dichotomy between *independent* and *dependent* entities (depending on the existence of other entities, such as quality, role, function); defines as *specifically dependent continuants* those that cannot migrate from one bearer to another; classifies *immaterial* entities in *sites*, *fiat boundary continuants*, and *spatial regions*, which are particularly important distinctions in geography and city planning; classifies *occurrent* entities in *processes*, *temporal regions, spatiotemporal regions;* and uses, as all ontologies, the distinction between *instances* (individuals, particulars) and *universals* (generals, types) (Fig. 5) [32].

The top-level ontology at the central hub created with BFO-ISO should contain all classes allowing the alignment of initial ontologies. The spatial regions and temporal regions should be defined in ways to allow matching with the spatial and temporal classes of initial ontologies, enabling spatial and temporal synchronization of data. New object and data properties should be defined at the level of the central hub across the initial ontologies.

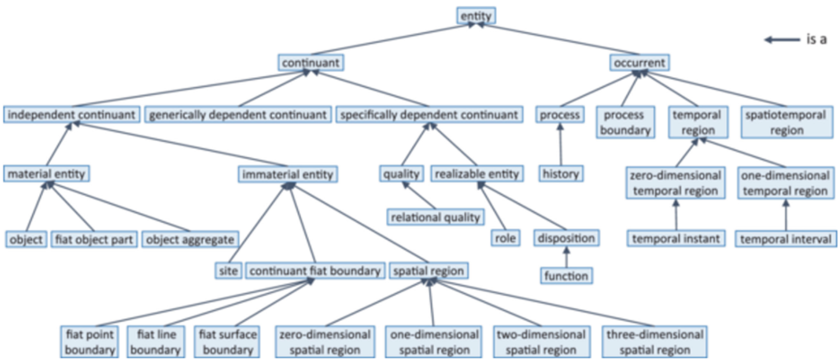

Fig. 5. The BFO-ISO ontology [33]

Transformations of Initial Datasets. Three types of transformations are needed to make city datasets semantically interoperable, semantic alignment, spatial alignment, and temporal alignment.

Semantic alignment is required if the same thing is denoted by different terms in different datasets, e.g., "innovation area" and "innovation zone" or “industrial district” and “industrial area”, or “yard” and “outdoor area”. In owl, “owl:equivalentClass” provides class equivalence, allowing a class description to have the same class extension as another class description, meaning that two classes are alternate names, are equivalent definitions of the same thing, or have the same set of instances.

Spatial alignment is required if data is recorded at different spatial entities, as shown in Fig. 2, at the building, building block, and city grid levels. If so, there is a need to introduce new spatial classes (fiat boundary or spatial region) that are common across datasets. Cities are full of physical boundaries and fiat boundaries that define land properties, administrative regions, postal districts, urban communities, and other human-induced demarcations. Data records follow these demarcations and to achieve data interoperability, there is a need to transform the initial datasets to new spatial categories common across datasets.

Temporal alignment is also required if data is recorded in different time periods. Introduction of common temporal entities with properties such as “hasBeginning” and “hasEnd”, and therefore common temporal intervals across datasets is a way to align datasets (see also, Time Ontology in OWL) [34]. There is also a need to transform the initial datasets to match the common temporal entities across datasets.

5 Use Case

An indicative use case scenario to highlight the potential use of the proposed framework in a SC which adopts all five axes is presented below;

Harbor authorities, recently included their legacy system, controlling the installed sensors network in the wider harbor area, into the proposed framework adapter. A cruise ship approaches the coastal city and the captain asks for available live sensors data (temporal axis) from the area around the port (spatial axis). Among various sensors streamed to the ship, the captain focuses on data coming from sensors measuring water height every 20 min (intensity axis). The selected interval is sufficient as the tide is relatively mild in this port. The extra information helped the captain to decide the most appropriate quay leading to faster debarkation. The harbor authorities have already agreed to gradually install a new set of similar sensors capable to stream water height on one-second interval, providing near-to real-time measurements allowing bigger ships to approach more safely. Although the streamlined data structure of the newly employed sensors is different, the framework is able to ingest and serve data in a unified way via the meta-ontology and enclosed adapter layer in the single point of access.

The passengers, during their trip, were wearing wristbands allowing them to unlock their cabins, purchasing goods, and benefiting cruise ship to optimize resources based on their movements (e.g., identifying understaffed areas on the deck). The wristbands are broadcasting a unique identifier but some elderly passengers also consent to wear more advanced ones that also monitor their medical condition (i.e., heart rate, oxygen saturation in the blood) and broadcast in real-time. These wristbands are compatible to the proposed framework and they are broadcasting anonymous data in hotspots around the city. In case of an unusual peak of heart pulse (variation-fluctuation axis) in one passenger's wristband, while navigating through the city, has triggered an alarm. Local health centers are constantly fetching anonymized health-related information about citizens via the framework adapter in a given area.

At a higher level, the local authorities are able to collect big data from the sensors around the city in a dedicated cloud-based application (edge capabilities axis) and process historical data, including cruise ship visits. Combining different sensor data (weather, traffic, pollution, electricity consumption, etc.) with static information such as rent prices, yearly events (e.g., conferences) they can get useful insights and decide on future policies that could benefit the city and its citizens.

The presented use case has demonstrated the applicability of the proposed framework by introducing interactions with each frameworks' axis. It is important to note that interoperability is achieved by the adapter layer in conjunction with the meta-ontology, the scenario that highlights this possibility is the integration of new water sensors more capable to measure in high frequency with legacy water sensors. Furthermore, the wristband scenario showcases the exploitation of combining multiple framework axis to drive operational efficiency and instantaneous response in case of emergency. Lastly, the local authority's scenario presents an envisaged monitoring setup that aggregates data across a variety of sensors with different edge computing capabilities across multiple smart city ecosystems.

6 Conclusion

In this article, we have introduced an IoT framework that has the capability of offering a comprehensive arrangement for interfacing with smart city IoT platforms and/or applications. The framework is structured around five distinct axes, each of which is focused on specific functionalities. Furthermore, the combination of these axes with different configurations has the potential to significantly enhance the information retrieval process. A top-level neutral ontology is introduced as an integration mechanism of domain-specific ontologies to reduce heterogeneity and establish a multi-axes alignment. Given the usability of each axis, a use case has been presented that highlights the utilizations and applicability in real-world scenarios. Finally, the adoption of the proposed framework by contemporary smart cities can support the transition to the meta-information era where big data demands efficient data management across a variety of data sources.

Acknowledgements. This research was carried out as part of the project "iWet: Intelligent IoT System for Quantitative and Qualitative Measurements for Water Distribution Networks" (Project Code: KMP6-0072250) under the framework of the Action "Investment Plans of Innovation" of the Operational Program "Central Macedonia 2014 2020" that is co-funded by the European Regional Development Fund and Greece.

References

1. Tsampoulatidis, I., Komninos, N., Syrmos, E., Bechtsis, D.: Universality and interoperability across smart city ecosystems. In: Streitz, N.A., Konomi, S. (eds.) Distributed, Ambient and Pervasive Interactions Smart Environments, Ecosystems, and Cities. Lecture Notes in Computer Science, vol. 13325, pp. 218–230. Springer, Heidelberg (2020)
2. Silva, B.N., Khan, M., Han, K.: Towards sustainable smart cities: a review of trends, architectures, components, and open challenges in smart cities. Sustain. Cities Soc. **38**, 697–713 (2018). https://doi.org/10.1016/j.scs.2018.01.053
3. Tang, S., Shelden, D.R., Eastman, C.M., Pishdad-Bozorgi, P., Gao, X.: A review of building information modeling (BIM) and the internet of things (IoT) devices integration: Present status and future trends. Autom. Constr. **101**, 127–139 (2019). https://doi.org/10.1016/j.autcon.2019.01.020
4. Costin, A., Eastman, C.: Need for interoperability to enable seamless information exchanges in smart and sustainable urban systems. J. Comput. Civ. Eng. **33**(3), 04019008 (2019). https://doi.org/10.1061/(asce)cp.1943-5487.0000824
5. Hassani, A., et al.: INFORM: a tool for classification and semantic annotation of IoT datastreams. In: 2021 IEEE 7th World Forum on Internet of Things (WF-IoT) (2021). https://doi.org/10.1109/wf-iot51360.2021.9594994
6. Haller, A., Janowicz, K., Cox, S., Le Phuoc, D., Taylor, K., Lefrançois, M.: Semantic Sensor Network Ontology (n.d.). https://www.w3.org/TR/vocab-ssn/. Accessed 20 Jan 2023
7. Gonzalez-Gil, P., Martinez, J.A., Skarmeta, A.F.: Lightweight data-security ontology for IoT. Sensors **20**(3), 801 (2020). https://doi.org/10.3390/s20030801
8. Ganzha, M., Paprzycki, M., Pawłowski, W., Szmeja, P., Wasielewska, K.: Semantic interoperability in the Internet of Things: an overview from the INTER-IoT perspective. J. Netw. Comput. Appl. **81**, 111–124 (2017). https://doi.org/10.1016/j.jnca.2016.08.007
9. Agarwal, R., et al.: Unified IoT ontology to enable interoperability and federation of testbeds. In: 2016 IEEE 3rd World Forum on Internet of Things (WF-IoT) (2016). https://doi.org/10.1109/wf-iot.2016.7845470
10. de Matos, E., et al.: Context information sharing for the Internet of Things: a survey. Comput. Netw. **166**, 106988 (2020). https://doi.org/10.1016/j.comnet.2019.106988
11. Syrmos, E., et al.: An intelligent modular water monitoring IoT system for real-time quantitative and qualitative measurements. Sustainability **15**(3), 2127 (2023). https://doi.org/10.3390/su15032127
12. Chong, H.F., Ng, D.: Development of IoT device for traffic management system. In: 2016 IEEE Student Conference on Research and Development (SCOReD) (2016). https://doi.org/10.1109/SCORED.2016.7810059
13. Avatefipour, O., Sadry, F.: Traffic management system using IoT technology - a comparative review. In: 2018 IEEE International Conference on Electro/Information Technology (EIT) (2018). https://doi.org/10.1109/eit.2018.8500246
14. Komninos, N., Bratsas, C., Kakderi, C., Tsarchopoulos, P.: Smart city ontologies: improving the effectiveness of smart city applications. J. Smart Cities **1**(1) (2016). https://doi.org/10.18063/jsc.2015.01.001
15. Rasmussen, M.H., Pauwels, P., Lefrançois, M., Schneider, G.F.: Building topology ontology (2021). https://w3c-lbd-cg.github.io/bot/, Accessed 20 Jan 2023
16. Kaebisch, S., McCool, M., Korkan, E., Kamiya, T., Charpenay, V., Kovatsch, M.: Web of Things (WoT) Thing Description 1.1 (2021). https://www.w3.org/TR/wot-thing-description/. Accessed 20 Jan 2023
17. Charpenay, V., Lefrançois, M., Villalón, M.P., Käbisch, S.: Thing Description (TD) Ontology (n.d.). https://www.w3.org/2019/wot/td. Accessed 21 Jan 2023

18. Katsumi, M.: Land use ontology (2023). https://enterpriseintegrationlab.github.io/icity/LandUse/doc/index-en.html. Accessed 23 Jan 2023
19. European Commission: Statistical classification of economic activities in the European Community NACE Rev. 2 (2006). https://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF. Accessed 24 Jan 2023
20. Risk Data Ontology - Open Risk Manual (2023). https://www.openriskmanual.org/wiki/Risk_Data_Ontology
21. Osman, I., Ben Yahia, S., Diallo, G.: Ontology integration: approaches and challenging issues. Information Fusion **71**, 38–63 (2021). https://doi.org/10.1016/j.inffus.2021.01.007
22. Ekaputra, F.J., Sabou, M., Serral, E., Kiesling, E., Biffl, S.: Ontology-based data integration in multi-disciplinary engineering environments: a review. Open J. Inf. Syst. **4**, 1–26 (2017)
23. Wang, J., Lu, J., Zhang, Y., Miao, Z., Zhou, B.: Integrating heterogeneous data source using ontology. J. Softw. **4**(8), 843–850 (2009). https://doi.org/10.4304/jsw.4.8.843-850
24. Dou, D., LePendu, P.: Ontology-based integration for relational databases. In: Proceedings of the 2006 ACM Symposium on Applied Computing (2006). https://doi.org/10.1145/1141277.1141387
25. Popplewell, K., Thoben, K.-D., Knothe, T., Poler, R. (eds.): Enterprise Interoperability VIII. PIC, vol. 9. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-13693-2
26. Ruttenberg, A., Smith, B.: JOWO-2019-Tutorial-BFO-ISO.pptx | Powered by Box (2019). https://buffalo.box.com/v/BFO-ISO-JOWO. Accessed 23 Jan 2023
27. ISO/IEC (2021). ISO/IEC 21838-1:2021. https://www.iso.org/standard/71954.html. Accessed 26 Jan 2023
28. ISO/IEC (2021b). ISO/IEC 21838-2:2021. https://www.iso.org/standard/74572.html. Accessed 26 Jan 2023
29. Smith, B., Ceusters, W.: Ontological realism: a methodology for coordinated evolution of scientific ontologies. Appl. Ontol. **5**(3–4), 139–188 (2010). https://doi.org/10.3233/AO-2010-0079
30. Hagedorn, T.J., Smith, B., Krishnamurty, S., Grosse, I.: Interoperability of disparate engineering domain ontologies using basic formal ontology. J. Eng. Des. **30**(10–12), 625–654 (2019). https://doi.org/10.1080/09544828.2019.1630805
31. Komninos, N., Panori, A., Kakderi, C.: The smart city ontology 2.0: assessing the components and interdependencies of city smartness (2021). www.preprints.org. https://doi.org/10.20944/preprints202108.0101.v1
32. Smith, B.: On classifying material entities in basic formal ontology. In: Interdisciplinary Ontology: Proceedings of the Third Interdisciplinary Ontology Meeting, pp. 1–13. Keio University Press (2012)
33. ISO/IEC (2023). ISO/IEC 21838-2:2021. https://www.iso.org/obp/ui/#iso:std:iso-iec:21838:-2:ed-1:v1:en. Accessed 26 Jan 2023
34. Cox, S., Little, C.: Time ontology in OWL (2022). https://www.w3.org/TR/owl-time/. Accessed 26 Jan 2023