

Article

Federated Data Modelling for Heritage Building Performance Management

Angelo Massafra ^{1,*}, Ugo Maria Coraglia ², Silvia Di Turi ³ and Domenico Palladino ⁴

¹ Department of Architecture, University of Bologna, 40136 Bologna, Italy

² Department of Engineering and Applied Sciences, University of Bergamo, 24044 Dalmine, Italy; ugomaria.coraglia@unibg.it

³ Energy Efficiency Unit Department, Italian National Agency for New Technologies, Energy and Sustainable Economic Development, 00123 Rome, Italy; silvia.dituri@enea.it

⁴ Energy Efficiency Unit Department, Italian National Agency for New Technologies, Energy and Sustainable Economic Development, 06124 Perugia, Italy; domenico.palladino@enea.it

* Correspondence: angelo.massafra@unibo.it

Abstract

The fragmentation of knowledge across multiple sources and players, coupled with limited access to information, is one of the main challenges for the performance-based management of heritage buildings. Despite recent research efforts in the field of Heritage Building Information Modelling (HBIM), this technology alone is insufficient for managing the variety of data related to building performance and is challenging for stakeholders without digital expertise to adopt. These limitations, along with advancements in knowledge technologies, have led to the emergence of federated data modelling approaches as a core strategy for managing the complexity of buildings' operational information. To address data fragmentation, this research proposes a methodology for linking heterogeneous data on historic building performance. The approach structures heterogeneous data, gathered from multiple sources—HBIM models, sensors, and energy bills—into knowledge graphs that enable semantic integration, cross-domain queries and support interactive visualisation. As part of the BeTwin research project, the methodology is validated through its application to a case study in the Appia Antica Archaeological Park (Rome, Italy).

Keywords: built heritage; building performance; decision support systems; building information modelling; knowledge graphs; digital twins

1. Introduction

Collecting reliable, systematised information to assess the energy performance of historic buildings is a fundamental step toward planning management strategies that improve efficiency without compromising heritage values. Nonetheless, obtaining accurate data in these contexts presents notable difficulties, particularly regarding the detailed characterisation of construction and mechanical systems.

Unlike modern constructions, heritage buildings are often characterised by unique and complex components, outdated construction techniques, and traditional materials whose thermal behaviour cannot be easily predicted using standard, contemporary theoretical models. Regarding the building envelope, for example, researchers highlight uncertainties in understanding thermophysical properties, owing to moisture and material

Academic Editor: Yung Yau

Received: 29 September 2025

Revised: 9 December 2025

Accepted: 17 December 2025

Published: 20 December 2025

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\) license](https://creativecommons.org/licenses/by/4.0/).

inhomogeneity [1]. The development of non-destructive methods, such as infrared thermography [2,3] and heat flux meters [4,5], led to the spread of tools capable of providing the critical parameters needed for energy diagnostics; however, these are not the only data required for performance characterisation. Beyond physical measurements, analysing historical documentation, archival data, and energy bills, as well as conducting geometric surveys and monitoring occupancy patterns and environmental conditions (e.g., through sensors), are crucial for building comprehensive knowledge bases.

Significant issues in the data collection process arise because, for these buildings, much of this information is difficult to obtain, time-consuming to gather, or, if already available, is spread among many actors in diverse formats, forming unconnected data silos [6]. Consequently, every time such data are needed from building managers (e.g., for conducting energy audits or planning renovation interventions), it must be requested from multiple actors, a process commonly carried out through several email or verbal requests that leads to information loss, frustration, and wasted time and resources during information exchange [7]. Additional problems occur because of the fact that performance data is not static and varies over time. This dynamism may not be compatible with the usual slowness of the data collection process.

In this perspective, two main gaps need to be filled. The first involves systematising performance data in digital environments so that building managers can access it quickly and efficiently, thereby using it to support their data-driven energy management decisions. Second, to build these digital systems, there is a need for knowledge modelling frameworks and technologies capable of hosting the different types of data characterising historic buildings and linking them within a systemic vision. These are two crucial prerequisites that, in addition to accelerating data collection, can lead to a shift in the performance-based management of historic buildings, here understood as the systematic process for monitoring, analysing, and evaluating building performance to align management practices with organisational goals in terms of efficiency and sustainability [8].

In response to these challenges, this paper presents a knowledge-modelling process that enables linking a set of building data collected from different sources in heterogeneous formats and forming robust knowledge bases for easy data querying and retrieval in support of building management processes. At the core of the methodology is a Federated Data Modelling (FDM) approach, a decentralised approach that creates virtual intermediate layers connecting disparate data systems, enabling seamless access, integration, and analysis of distributed information as if it were in a single, unified database [9].

To operationalise the FDM methodology, the research adopts information technologies such as HBIM and Knowledge Graphs (KGs), leveraging the BTwin toolkit, a low-code, open-source solution for prototyping web-based decision support systems (DSSs) for building management developed at the University of Bologna [10]. The method is applied in a real environment—the Capo di Bove building complex, a representative case study drawn from the Parco Archeologico dell'Appia Antica (PAAA), located in Rome (Italy).

In previous work, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA) compiled a comprehensive dataset for this complex, combining geometric surveys, non-destructive tests, in situ investigations, environmental monitoring, and energy bills [11]. This dataset supported the assessment of envelope thermal behaviour, indoor environmental conditions, and operational performance, providing a basis for energy audits. In the present study, these data are structured within a Knowledge Graph (KG) using a Federated Ontology Framework (FOF) to feed a user-friendly web application that visualises key performance indicators through intuitive dashboards. The outcome is a prototype platform that enables managers to extract relevant information for performance-based management easily.

2. Background and Related Works

2.1. Information Modelling for Heritage Buildings

The fragmentation of data among multiple stakeholders and within isolated silos is one of the most significant challenges to the digitalisation of the built heritage field.

This issue has been addressed in recent years by many research efforts, aiming to boost the digital transition of the sector, which have proposed a suite of frameworks, tools, and methods to harmonise heterogeneous data sources and converge data collection campaigns into unified digital systems [12]. Such systems, briefly described in this section, allow the production of information models or, more generally, databases to capture the geometric composition of heritage asset elements and enrich them with relevant information, for instance, concerning performance characteristics.

One of the most well-known technologies is HBIM, a methodology introduced by Murphy et al. in 2009 [13] that, due to its high potential and despite its substantial limitations, is still in development nowadays [14]. In contrast to the BIM methodology, which usually focuses on the design and construction of new buildings, the addition of the letter H (“historic” or “heritage”) to the acronym shifts the attention to the digital representation of the issues characterising heritage buildings and the procedures for managing and safeguarding them. HBIM therefore adapts all the methodological aspects of BIM, which aim to create digital models for managing data throughout the entire life cycle of a building, to more specific building components of built cultural heritage [14,15].

Despite its theoretical capability to promote standardised approaches for managing historic buildings’ data and sharing knowledge between the different actors involved in building management [16], HBIM still presents several bottlenecks that hinder its widespread adoption in professional practice. On the one hand, current commercial BIM software has inherent limitations because it is primarily co-designed for new construction. This makes it difficult to standardise the semantics of heritage structures and to model the customised geometries of historic assets in 3D accurately [14]. On the other hand, HBIM—and BIM in general—is not intended for managing dynamic or time-series data. Therefore, to enable continuous performance monitoring over time, it needs to be integrated with other technologies within more articulated digital systems, such as Digital Twins (DTs) [17].

2.2. Knowledge Technologies for Data Integration in AECO

To face the data fragmentation issue, recent years have seen a growing interest in semantic web technologies and Linked Data (LD), which, initially developed in the IT domain, have begun to be adopted within the Architecture, Engineering, Construction, and Operations (AECO) [18]. Researchers have indeed recognised that LD’s adaptable and scalable nature can facilitate the incorporation of varied information sources, such as BIM and HBIM, as well as the Internet of Things (IoT) and Geographic Information Systems (GIS), without being confined to a static data structure [19].

In particular, graphs have emerged as a core technology for federating cross-domain data in the construction industry, enabling interoperability among available sources [20,21]. These are data structures that represent the entities of a given domain and their interconnections as networks composed of nodes, corresponding to objects, and edges, corresponding to the relationships between them. In such structures, both nodes and edges can store attributes and properties (as in Property Graphs, PGs), semantic labels (as in Labelled Property Graphs, LPGs), or ontological references (as in Knowledge Graphs, KGs) [22].

The inherent structure of KGs allows data to be linked to ontological concepts, which unambiguously define their semantic meaning and provide contextual information. In

this way, KGs facilitate a structured yet flexible representation of a domain in accordance with LD principles, thereby fostering interoperability among heterogeneous data environments [23]. For this reason, they are ideally suited to operationalising the FDM approach within the AECO context. In addition, compared to more traditional structures like relational databases, graph database engines typically use storage structures that are more efficient for managing interconnected data. This architecture reduces the computational cost of queries involving transversal connections, thereby speeding data retrieval. As a result, in modern DSSs, graph-based models often constitute the core layer for data retrieval, supporting the direct development of visualisation services on top of them.

2.3. Federated Data Modelling for Building Performance Management

In the FDM approach, ontologies are essential to attribute semantic and contextual meanings to the KG. As the core semantic structure of KGs, ontologies provide formal, shared descriptions of concepts, properties, and relationships within a given domain, enabling the adoption of a common vocabulary that ensures consistent structure, logical reasoning, and unambiguous data exchange [24].

Given their fundamental role and the adoption of knowledge technologies, AECO research has focused on developing ontologies for multiple purposes, spanning from building design to performance management. This development has led to several scientific initiatives in the field. For instance, the Built Environment Ontology Lookup Service was created to support practitioners with a publicly accessible, expert-validated online repository of AECO ontologies [25].

A literature review by Pritoni et al. [25] recently examined the wide range of ontologies supporting building energy applications. In this paper, we limit ourselves to presenting only those relevant to the methodological section (Section 3) (i.e., IFC, BOT, Brick, SSN, SOSA, and EM-KPI), leaving the reader to investigate the topic further through the referenced work.

IFC (Industry Foundation Classes) is the open standard for BIM; maintained by buildingSMART International, it provides standardised, vendor-independent and machine-interpretable definitions for all aspects related to the physical and functional properties of a building [26]. BOT (Building Topology Ontology), instead, is an ontology developed by the W3C Linked Building Data Community Group [27]; it provides a concise vocabulary for defining the fundamental concepts related to buildings' spatial elements and their topological relationships. Brick is an open-source metadata schema tailored to smart buildings; it standardises the representation of physical, logical, and system assets (e.g., HVAC zones, sensors) and their interconnections for operational analytics [28]. While SSN (Semantic Sensor Network Ontology) and its simplification, SOSA (Sensor, Observation, Sample, and Actuator), are ontologies that model sensors, their observations, deployments, procedures, and actuators, supporting semantic annotation of building-automation data [29,30]. Finally, EM-KPI (Energy Management Key Performance Indicator Ontology) is an ontology that defines a structured vocabulary for energy-performance indicators at building and district scales, facilitating reporting, benchmarking, and analysis of energy management data [31].

Each of these ontologies focuses on one or more subdomains of buildings, but cannot cover the entire domain through an "all-encompassing" approach. Therefore, they are usually combined under FDM frameworks to provide cross-domain digital representation.

For example, Merino et al. [6] leveraged both IFC and Brick within a KG framework to create a DT of an educational building for HVAC fault detection and diagnosis in FM. Similarly, Chamari et al. [32] explored the integration of building management systems, IoT, and BIM using RDF (Resource Description Framework) graphs based on IFC and

Brick schemas, and proposed a hybrid workflow for handling time-series sensor data. Ramonell et al., working with Neo4j, a well-known graph database engine, proposed an LPG architecture that semantically maps multiple ontologies, positioning the graph as the interoperability backbone of their digital twin system [21].

Applications in the heritage field were also developed. For instance, Niccolucci and Felicetti [33] presented the Heritage Digital Twin Ontology, where HBIM-derived structural and material information feeds into a semantic graph that integrates real-time sensor streams. Hosamo and Mazzetto proposed an ontology-based KG for heritage building conservation purposes [34]. Or, Falcone et al. [35] proposed a graph-based data management workflow that integrates HBIM-derived 3D models with a property graph for continuous monitoring of cultural heritage assets.

All these contributions underscore the importance of KG as an ontological and semantic core layer for data interoperability within DSS.

2.4. Research Context

To address data fragmentation in historic building performance characterisation processes, this research introduces an FDM approach to link building-related information from multiple sources, stored in heterogeneous formats. The method consolidates fragmented data into a robust knowledge bases that enable efficient data querying and retrieval to support building management processes.

The methodology is validated through a significant case study, which serves as the foundation for a prototype DSS that provides building managers with user-friendly dashboards to facilitate access to KPIs. Although the methodology is here applied to energy performance data, the proposed approach is adaptable to other categories of performance-related information within the built heritage domain.

This work falls under the BeTwin research project at the University of Bologna (“Building Digital Twins for Built-Heritage Performance-Based Management”), advancing the foundations presented in previous research [10,36,37]. It was conducted through a collaboration with ENEA and PAAA, which, within the activities of the Electric System Research Plan 2022–2024 [11], collected a vast amount of data about some buildings located in the park.

In particular, within the PAAA’s portfolio, the Capo di Bove complex was selected as a representative demonstration site. This choice was motivated by several factors, including its relevance as a medium-sized protected historic complex with forthcoming energy renovations, its suitability for testing non-invasive and potentially energy-retrofit solutions, and the opportunity to collaborate with building administrators during data collection.

3. Methodology

3.1. Workflow Overview

In this section, we describe the FDM methodology used in the research, along with the tools employed. Figure 1 depicts the workflow followed for delivering the federated KG and the connected prototype DSS. It consists of the following phases:

- Data acquisition: collection and acquisition of the primary data;
- Data modelling: realisation of the HBIM model and preprocessing of the time-series;
- Data processing: aggregation of the raw data and calculation of relevant KPIs;
- Data integration: data linking via the federated KG;
- Data visualisation: visualisation of performance data in interactive dashboards.

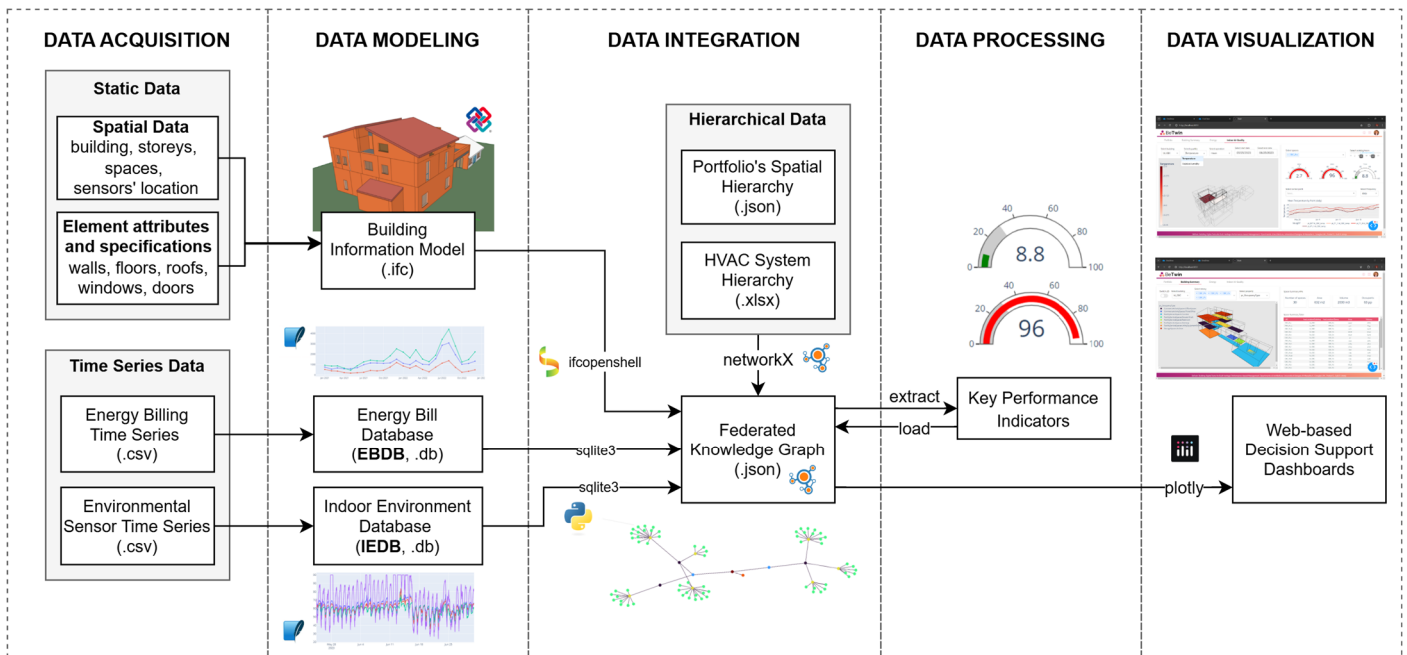


Figure 1. Methodological workflow for delivering the federated KG and the prototype DSS.

3.2. Data Acquisition Phase

For the chosen case study, the data acquisition process, performed by ENEA, was oriented toward collecting the data needed to conduct an energy audit, using the ISO 52016 method [38].

First, archival research was conducted to gather all relevant data on the buildings, their contexts, and their chronological features. Direct surveys were conducted thanks to the support of the PAAA’s staff, including metric measurements, photographic documentation, site visits, and non-destructive investigations, such as infrared thermography and heat flux measurements, to, respectively, identify thermal bridges, material discontinuities, and stratigraphic differences, and estimate the actual thermal resistance of the key opaque envelope components. These investigations enable a detailed characterisation of the envelope, resulting in a CAD inventory of opaque and glazed components with geometric, construction, and thermophysical properties (where possible, measured; otherwise, deduced). After this, data on the main components of the mechanical and electrical systems were collected and summarised in an XLSX inventory. Electricity utility bills were also gathered. Finally, using temperature and humidity sensors, microclimatic monitoring was conducted in several main spaces to understand indoor environmental conditions.

At the end of the campaign, both static and dynamic data were available. The static data concerned the spatial organisation of the selected site, as well as the material, construction, age, and thermal-physical properties of the building elements. The dynamic data consisted of CSV time series containing monthly utility billing records and hourly environmental monitoring data.

3.3. Data Modelling Phase

The first action in the data modelling phase was the delivery of HBIM models, which were realised through Autodesk Revit 2023, using the data acquired in the previous phase.

The main aim of the HBIM process was to produce information models containing relevant data on the spatial layout of buildings, the material characteristics of envelope and partition components, and the positioning of sensors. The relevant data about construction components was taken from the envelope element inventory furnished by ENEA (also reported in the 2023 extended report entitled “Energy efficiency of historic and listed

building complexes: characterisation of real case studies (LA1.4)" [11]) and inserted within element types. Table 1 and Figure 2 depict an example of significant material data for a masonry type.

Several constraints arose during the modelling. For instance, in some cases, given the stratigraphic complexity, the masonry was simplified into homogeneous sections, maintaining the prevailing stratigraphy as a reference. Moreover, modelling has required the definition of customised parameters. These were managed using Revit's parameter functions for "Window", "Wall", "Floor", "Door", and "Roof" object categories (see Figure 2).

Table 1. Example of table relating to the Conference room Masonry (CB-M1) [11].

Layer	Description	Thickness (m)	Thermal Conductivity (W/mK)	Thermal Resistance (m ² K/W)	Density (Kg/m ³)	Specific Heat (J/KgK)	Thermal Transmittance (W/m ² K)
	Rsi	-	-	0.1299	-	-	-
1	Internal Plaster	0.02	0.7	0.029	1400	1000	-
2	Bricks	0.05	0.36	0.137	600	840	-
3	Air gap	0.06		0.18	1.33	1008	-
4	Tuff blocks	0.6	0.7	0.86	1600	1000	-
	Rse	-	-	0.04	-	-	-
		0.73	1.373	1.376	-	-	0.727

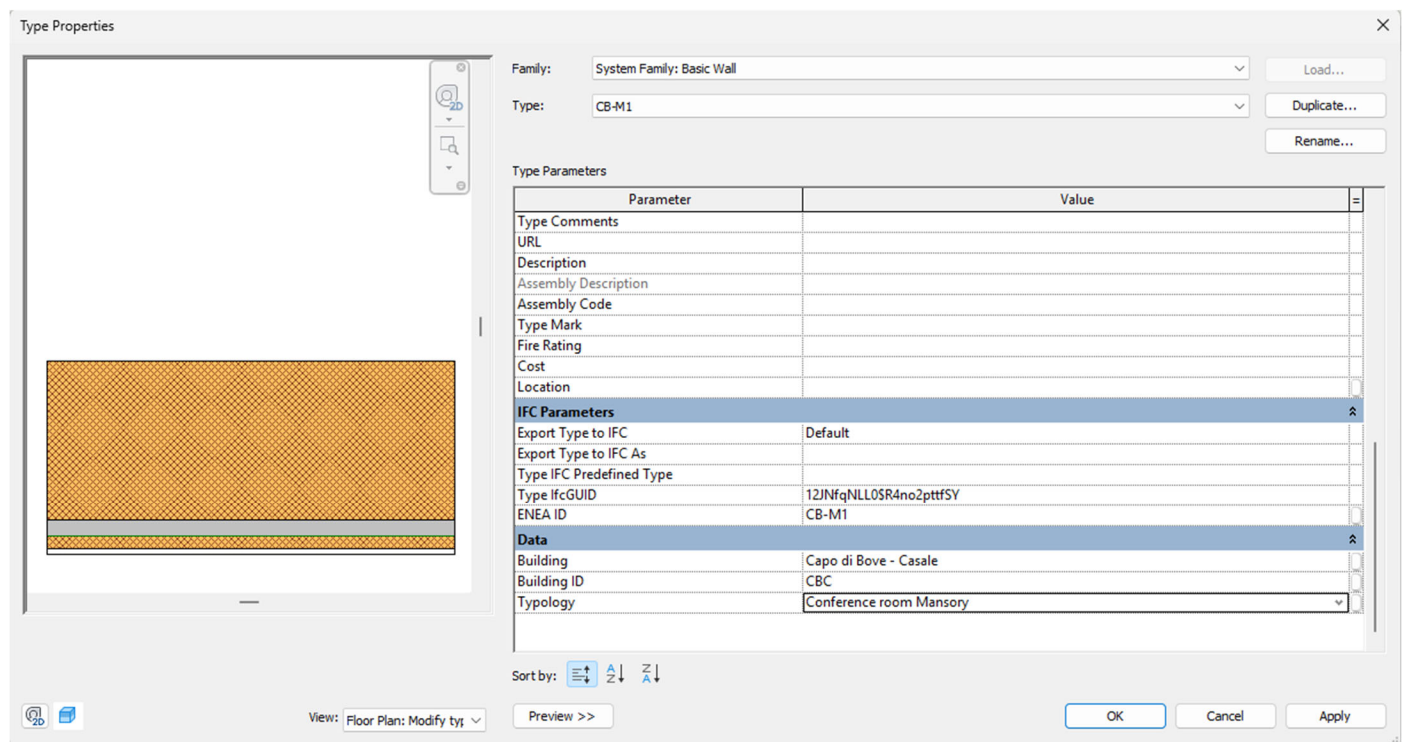


Figure 2. Custom parameters according to project requirements.

After producing the HBIM, time-series data were prepared for ingestion into the DSS.

To maximise computational performance, billing records and sensor readings were stored in two different dedicated SQLite databases: the Energy Bill Database (EBDB) and the Indoor Environment Database (IEDB). These time-series databases utilise a minimal schema comprising the sensor's unique identifier, the timestamp, and the measured value and its corresponding unit, enabling rapid sequential writes, efficient indexing, and lightweight storage.

Referential integrity between time-series observations and the element metadata stored in the HBIM was maintained by linking each record to its corresponding element via the element identifier. In particular, each observation in the EBDP was linked to an energy meter ID, and each observation in the IEDB was linked to a sensor ID.

3.4. Data Integration Phase

The data integration phase was articulated according to three main subphases, which are:

- (1) the establishment of an FOF;
- (2) the serialisation of data;
- (3) the generation of the KG.

3.4.1. Federated Ontology Framework

To enable data integration, an ontology federation process was implemented to identify, map, and interlink all the entities and relationships that define the building performance knowledge domain.

This process resulted in the creation of a Federated Ontology Framework (FOF), a high-level semantic layer designed to connect and coordinate multiple existing ontologies, each developed independently and focused on a specific domain aspect (e.g., topology, sensors, systems, etc.). The FOF acts as a bridge between many subontologies by establishing correspondences between equivalent or related concepts and defining how data from one ontology can be interpreted in relation to another.

Federation was chosen to avoid creating a new, monolithic, all-encompassing ontology for building performance; an approach that would have duplicated efforts and increased redundancy in a domain already rich with specialised semantic systems [39]. Instead, the FOF leveraged existing ontologies—i.e., IFC, BOT, Brick, SSN/SOSA, and EM-KPI—to promote reusability.

The FOF, depicted in Figure 3 and described below, was encapsulated within the BTwin Python library (vv.0.4.0) [10], which was used to operationalise the FDM method and the ontology within a KG. The FOF is composed of different modules (i.e., the “Document” module, the “Equipment” module, the “Spatial Element” module, the “KPI” module, and the “Interface” module) that harmonise the aforementioned standards.

At the core of the FOF, BOT allows for representing buildings’ spatial hierarchy with classes such as “bot:Building”, “bot:Site”, “bot:Storey”, and “bot:Space”. This hierarchy is supported by the “brick:hasLocation” relationship, which enables the placement of spatial elements, equipment, and sensor objects within higher-level spatial elements. The Topologic class hierarchy [40] complements this structure, allowing for modelling the elements that bound spaces, including walls, floors, and roofs (“top:Face” class), as well as windows and doors (“top:Aperture” class). IfcOWL (IFC Ontology Web Language) maps element attributes into “ifc:Property” entries, grouped within “ifc:PropertySet”, and allows representing physical sensor devices, located in the HBIM, through the “ifc:Sensor” class. The Brick schema refines sensor modelling by classifying sensing devices as “brick:Sensor” and their individual measurement channels as “brick:Point”, linking them together by the “brick:hasPoint” relationship; Brick also connects sensors to spaces using “brick:hasLocation” and groups rooms into larger zones (e.g., thermal zones) via “brick:Zone”, thereby interoperating with BOT’s spatial hierarchy. To capture sensor and meters observations, the SSN and SOSA ontologies introduce “ssn:Observation” and “sosa:ObservableProperty” classes. Observations are linked to their generating sensors through “sosa:madeBySensor” and grounded in temporal concepts (“time:Instant” and “time:Interval”) to record when data are collected. The EM-KPI ontology contributes with classes such as “eko:KPI”, “eko:KPICalculation”, and “eko:KPIValue” that allow for

defining and storing performance metrics; these are tied back to spatial elements via “eko:hasAssociatedObject”, with temporal classes again used to trace metric evolution over time. Finally, we introduced a custom class, “btwin:Document”, to map all the digital artefacts (databases, documents, and models) to the main spatial elements. As described in Section 4, for this application, for instance, each “bot:Building” was linked via “btwin:hasDocument” to its IFC model and to the endpoints of the corresponding EBDB and IEDB, all representing “btwin:Documents”.

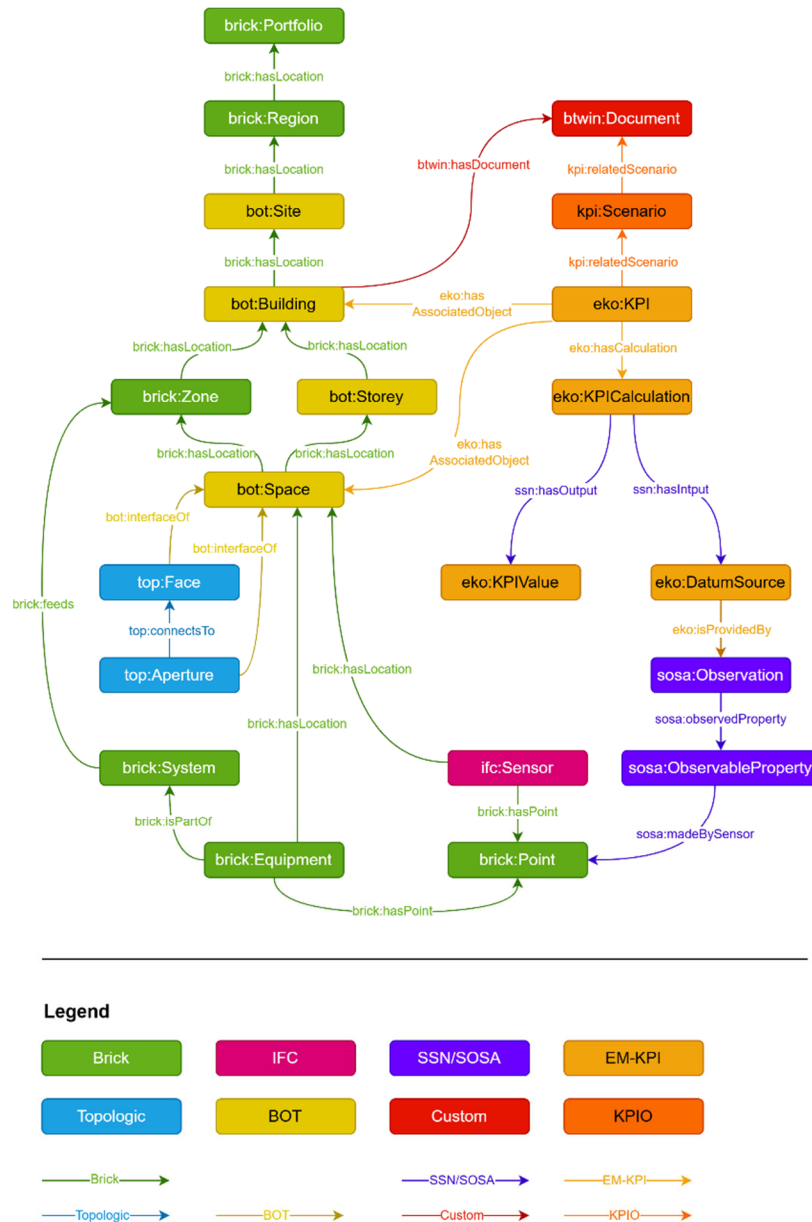


Figure 3. Federated ontological framework supporting the data integration process.

The FOF therefore serves as an intermediate layer for aligning the data schemas of external sources (i.e., the HBIMs, the SQLite DBs, and the XLSX inventory spreadsheets) to a more coherent semantic logic.

3.4.2. Graph-Based Data Serialisation

The FOF was serialised using JSON-LD (JavaScript Object Notation for Linked Data), a lightweight format for creating machine-readable data, optimised for the web. JSON-LD represents graphs as lists of JSON objects, where each object contains a unique identifier

("@id"), a declaration of its ontological class ("@type"), and key-value mappings that encode its relationships to other objects in the graph.

To translate native data into this JSON-LD schema, a suite of connectors from the BTwin toolkit was used; each connector transforms schema-specific constructs into ontology-compliant JSON entities. For example, BTwin's IFC Connector, built on IfcOpenShell, was used to parse IFC files, extract IfcBuilding, IfcBuildingStorey, and IfcSpace objects along with their properties, and emit corresponding JSON objects (Figure 4). The SQLite connector, in contrast, was used to map sensor points and SSN observations to the corresponding sensor devices and the spaces that host them.

JSON-LD

```
{
  "@context": {
    "brick": "https://brickschema.org/schema/1.3/Brick",
    "bot": "http://w3id.org/bot",
    "ifc": "https://standards.buildingsmart.org/IFC/DEV/IFC4/ADD2_TC1/OWL"
  },
  "@graph": [
    {
      "@id": "bl_CBC",
      "@type": "bot:Building",
      "relationships": { "brick:hasLocation": [{"@id": "st_CBPBV", "@type": "bot:Site"}] },
      "name": "Main Building - Capo di Bove Site"
    },
    {
      "@id": "CBC_PT",
      "@type": "bot:Storey",
      "relationships": { "brick:hasLocation": [{"@id": "bl_CBC", "@type": "bot:Building"}] },
      "name": "Ground Floor"
    },
    {
      "@id": "CBC_PT.1",
      "@type": "bot:Space",
      "relationships": { "brick:hasLocation": [{"@id": "CBC_PT", "@type": "bot:Storey"}] },
      "name": "Conference Room"
    }
  ]
}
```

IFC (EXPRESS)

```
#100=IFCBUILDING('3$gs5S34XC4fi69G$dtmE8',#20,'bl_CBC',,$,$,#16,$,'Main Building - Capo di Bove',.ELEMENT.,,$,#99);
#101=IFCAXIS2PLACEMENT3D(#3,$,$);
#102=IFCLOCALPLACEMENT('16,#101);
#103=IFCBUILDINGSTOREY('0efE_ZRvD8Uuv$PG7CQA_j',#20,'CBC_PT',$,'Level:Level',#102,$,'CBC_P0',.ELEMENT.,0.);
#104=IFCCARTESIANPOINT(0.,0.,429.99995279999996);
#105=IFCAXIS2PLACEMENT3D(#104,$,$);
#106=IFCLOCALPLACEMENT(#16,#105);
#5171=IFCSPACE('3qrPS7HbJ4MRxnWjnGrU4J',#20,'1829',,$,$,#5039,#5170,'Conference Room',.ELEMENT.,.SPACE.,,$);
#5172=IFCPACETYPE('2mKLiZifX11xItzU9B9fMJ',#20,'Space 1829:508270',,$,$,$,'508270',,$,.NOTDEFINED.,,$);
#5222=IFCPROPERTYSINGLEVALUE('UID',$,IFCTEXT('CBC_PT.1'),$);
#5223=IFCPROPERTYSET('0I3HjWgwv61PFw2dKutj3v',#20,'Altro',,$,($295,#296,#5219,#5220,#5221,#5222));
#5224=IFCRELDEFINESBYPROPERTIES('3dtQFG1Zr798PEZF5jCj11',#20,$,$,($5171),#5223);
#111655=IFCRELAGGREGATES('0efE_ZRvD8Uuv$PGFCQA_j',#20,$,$,#103,($1975,#2531,#2617,#2945,#3918,#4118,#5171,#5399,#6157,#6412,#6788,#7118,#7176,#7577));
```

Figure 4. Example of alignment between a Brick/BOT JSON-LD model and an IFC4 (EXPRESS) model for the main building of Capo di Bove complex: the building (bl_CBC), its ground floor storey (CBC_PT), and the Conference Room space (CBC_PT.1) are represented, with the IFC IFCSPACE entity carrying the same UID (CBC_PT.1) via an IfcPropertySingleValue.

3.4.3. Generating the KG

Once all data were serialised in JSON-LD, the resulting JSON file was processed using BTwin methods, which, when given this file as input, instantiate a Neo4j graph object. In this step, each entity and relationship defined in JSON was translated into nodes and edges of the Neo4j graph, which acts as an intermediate abstraction layer that mirrors the semantic connections of the federated data model in a computationally efficient form.

Importantly, the approach does not duplicate or centralise the original datasets. Instead, the primary data remain stored within their native environments—such as the HBIM models, and the time-series databases—and are accessed on demand through their semantic links. This design preserves data integrity, ownership, and traceability, while still offering the flexibility and interoperability required for integrated and targeted building performance analyses.

3.5. Data Processing

At this stage of the workflow, the graph and complementary time-series databases were jointly interrogated to derive meaningful building performance indicators.

Specifically, the EBDB was queried to compute three primary KPIs at the bot:Building level.

- Energy Usage (EU): Calculated by summing all interval readings over the period of interest.
- Energy Cost (EC): Obtained by applying the related tariff rate to the energy consumption derived from energy bills.
- Equivalent CO₂ Emissions (CO₂E): Derived by multiplying the energy usage by an appropriate emissions factor.

These KPIs were aggregated over monthly, yearly, and custom-defined periods, and normalised based on the building's net floor area and volume.

In addition, the IEDB, including indoor microclimatic conditions obtained from in situ measurements, namely, interior and exterior temperatures and relative humidity, was processed to aggregate microclimatic data at bot:Space level. Such aggregation was performed at hourly resolution, enabling the calculation of meaningful indicators, such as the average, minimum, and maximum temperature and humidity, over the entire monitoring period. With further development of the research, this data will serve as the foundation for a more comprehensive Indoor Environmental Quality (IEQ) database, which will be designed to support thermal comfort analysis.

All KPIs, once calculated, were serialised and added to the Neo4j graphs as nodes classified under the "eko:KPI" class. Moreover, they were linked back to the corresponding spatial elements ("bot:Building" and "bot:Space" nodes) via the "eko:hasAssociatedObject". Through this graph representation, users can perform cross-domain queries—from high-level building metrics to space-level metrics and to specific meter or sensor observations, and vice versa—exploring relationships and dependencies among entities originating from different source models, or applying selective extraction of relevant subsets of information for analysis or visualisation.

3.6. Data Visualisation Phase

Once linked in the KG, data were extracted from Neo4j using Cypher queries and directly fed into a set of web dashboards for data visualisation, which were built using Plotly (vv. 6.2.0) and Dash (vv. 3.3.0). These dashboards, developed according to Business Intelligence (BI) principles, included:

- A section where the model's essential geometry is viewed in 3D, with spaces colour-coded according to their properties or KPIs;
- A section displaying building-level KPIs, via gauge charts and other types of visualisations;
- A section showing time-series plots for more in-depth analysis of the raw data stored in the SQLite databases, applicable for inspecting the metrics of the building and its most significant spaces.

Each of these sections is equipped with dropdown menus and other controls to filter the data according to user requests, e.g., selecting a specific room or floor, choosing the property used to colour the geometry, picking which KPIs to display, or applying temporal and typological filters to the time series.

4. Demonstration

4.1. Case Study Description

The selected case study is the Capo di Bove complex, a historic site located on the Via Appia Antica in Rome. Managed by PAAA, the site extends over 8600 m² area comprises three main parts: an archaeological area, a principal three-level building, and a single-story outbuilding (Figure 5).

The site includes several volumes constructed in different historical periods and with varied construction techniques. Originally erected atop the remains of a Roman cistern, it underwent substantial modifications between 1949 and 1965, when it was transformed into a private villa. During this renovation, the external walls were built using *spolia* (ancient Roman and medieval architectural and sculptural fragments), including bricks and amphorae.

Currently, the main building hosts a range of cultural and administrative functions, including PAAA's headquarters, archives, and public areas for exhibitions, events, and educational activities. The outbuilding works as a visitor reception facility with restrooms and vending machines.

For ease of reading, technical details on the setting and collected data are reported in Appendix A: building envelope (Appendix A.1), MEP systems (Appendix A.2), energy bills (Appendix A.3), and microclimatic conditions (Appendix A.4). The results of applying the methodology to the case study are instead presented in the following text.



Figure 5. The Capo di Bove complex: main building, on the left, and outbuilding, on the right.

4.2. HBIM Models

The HBIM models shown in Figure 6 were created in Autodesk Revit 2023 using geometric and alphanumeric information derived from the data collection phase, summarised in Appendix A.

According to the project's goals, the model was developed at Level of Development (LOD) 300, with reduced geometric complexity, to accelerate digitalisation while preserving interoperability. Indeed, rather than produce a detailed geometric virtual reconstruction of the buildings, the HBIM objective here was to build a "skeleton" of elements to which performance data could be attached.

The models were exported in IFC format. In particular, rooms were exported as IfcSpace entities. Before export, these elements were enriched with functional and environmental attributes, including for each space a unique identifier (UID), the net area and volume, the designated function (classified according to OmniClass Table 13 [41]), the

people count and other relevant occupancy requirements, the main thermal setpoints in winter and summer, and basic HVAC characteristics (i.e., whether the space is heated, conditioned, and/or mechanically ventilated).

Opaque components, such as walls, floors, and roofs, were modelled using Revit's material layer sets and exported as IfcWall and IfcSlab entities. The materials composing the layer sets were associated with thermophysical properties, including density, thickness, thermal conductivity, and specific heat capacity. Then, from these values, Revit automatically derived aggregate parameters, including overall thickness, mass per unit area, thermal resistance, and thermal transmittance. A typology-based approach was adopted through a structured schedule of construction types.

Doors and windows were also modelled following a typological approach, using Revit families and ad hoc types developed for the case study, and exported as IfcDoor and IfcWindow entities. Each type was associated with essential thermal properties, including transmittance values for the glazing, frame, and combined unit, as well as the frame material and glazing type.

Sensors were also modelled in the HBIM using placeholder objects, each with an assigned UID matching that in the Neo4j graph, and then exported as IfcSensors. These placeholders were linked to corresponding spaces via containment relationships, enabling automated linkage to time-series sensor data during the integration phase.

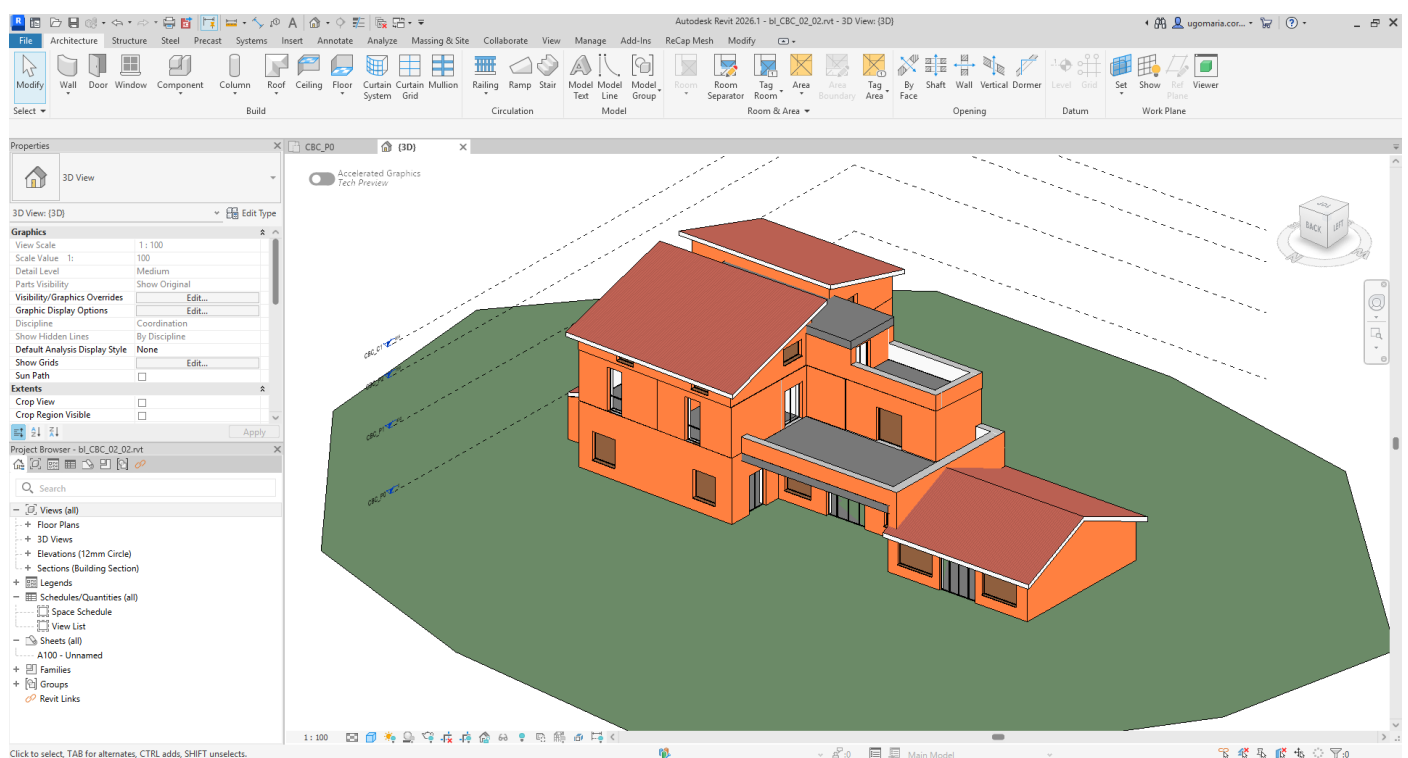


Figure 6. HBIM model of the case study in Autodesk Revit.

4.3. MEP System Inventory

Due to the difficulty of accurately tracking all MEP (Mechanical, Electrical, and Plumbing) components in the case study—limited access to physical installations, incomplete documentation, and the high level of modelling detail required to capture complex, intertwined service networks—the MEP systems were not modelled directly in the BIM environment. Indeed, attempting to model all these elements geometrically would have introduced excessive complexity and computational load without offering proportional analytical benefits.

Instead, the MEP systems were documented in an external spreadsheet (XLSX format), which served as a structured, lightweight system inventory (Table 2). This tabular format allowed the MEP data to be systematically described, capturing only essential information such as component type, producer, model, and spatial location.

This dataset was imported into the Neo4j graph database, where each spreadsheet row, corresponding to an individual equipment device, was represented as a Neo4j node. Each node included a reference to its ontological class (specified in the “type” column) and its connections to the building’s systems and spaces, thus preserving semantic relationships without requiring geometric modelling.

Although the geometric representation of MEP components was omitted, this approach ensures future interoperability: the same referencing strategy used for sensors can later be applied to link MEP records to BIM objects, should geometric or spatial data become available.

Table 2. Extract from the MEP spreadsheet.

ID	Name	Type	brick:isPartOf System	brick:hasLocation Space
AHU001	AHU (below ceiling)	brick:Air_Handling_Unit	heatingSystem	CBC_PT.1
AHU002	AHU (below ceiling)	brick:Air_Handling_Unit	heatingSystem	CBC_PT.1
FC001	Fan coil	brick:Fan_Coil:Unit	heatingSystem	CBC_PT.3
HP001	Heat pump	brick:Packaged_Heat_Pump	heatingSystem	CBC_Ext
BL1.4-1	How water boiler	Brick:Electric_Boiler	dhwSystem	CBC_P1.4

4.4. Timeseries Databases

Two distinct time-series repositories were created to manage dynamic data: the EBDB and the IEDB.

The EBDB stored monthly electricity consumption (in kWh) and associated costs (in euros) for each of the two buildings on the site —the main building and the outbuilding —from January 2021 to December 2022. Table 3 presents an excerpt from the database in tabular form. Each entry includes a unique ID (“Observation UID”), the meter that recorded the data (“sosa:madeBySensor”), the type of measurement (“Quantity”), the unit (“Unit”), the actual value (“Value”), and the date of observation (“Timestamp”).

The sensor repository, i.e., the IEDB, offers higher-resolution temperature and relative humidity data collected in three representative indoor spaces within the main building and at one outdoor reference location. Table 4 presents an excerpt from this database in tabular form, organised as previously to maintain consistency in dynamic data mapping.

Table 3. Extract from the time-series database with bill recordings.

Observation UID	sosa:madeBySensor	Quantity	Unit	Value	Timestamp
4113968597	pt_CBC_EEU	ElectricEnergyUsage	kWh	4552.00	2021-01-01
4121189850	pt_CBC_EEU	ElectricEnergyUsage	kWh	4156.00	2021-02-01
4129153564	pt_CBC_EEU	ElectricEnergyUsage	kWh	4242.00	2021-03-01
4136520893	pt_CBC_EEU	ElectricEnergyUsage	kWh	3424.00	2021-04-01

Table 4. Extract from the time-series database with sensor recordings.

Observation UID	sosa:madeBySensor	Quantity	Unit	Value	Timestamp
temp1	pt_PT_1-bl_CBC_temp	Temperature	C	22.00	24/5/23 10.00
temp2	pt_PT_1-bl_CBC_temp	Temperature	C	22.04	24/5/23 10.10
temp3	pt_PT_1-bl_CBC_temp	Temperature	C	22.07	24/5/23 10.20
temp4	pt_PT_1-bl_CBC_temp	Temperature	C	22.09	24/5/23 10.30

4.5. Knowledge Graph

After producing the HBIM and systematising the time-series, a Neo4j graph was created, following the workflow introduced in Section 3, which transferred all project data into a unified, federated, and semantic representation suitable for data-driven performance analysis.

A root “bot:Site” node, representing the Capo di Bove complex, was first instantiated to provide geographical context. The “Building–Storey–Space” hierarchy was then obtained via the IFC-to-Neo4j connector, which preserved native IFC spatial relationships and mapped them into the KG. The “bot:Building” nodes were subsequently linked to the site.

Sensors embedded in the IFC were also ingested through the IFC connector. Their spatial containment was defined in Python using Topologicpy [42], which ensured that each sensor node was correctly assigned to its host space.

Separate point entities representing the individual measurement channels (temperature, humidity, electricity consumption, and cost) were instead created manually and mapped to the appropriate Brick classes.

MEP system and device nodes were generated by a converter that translated the XLSX equipment schedule into Brick-aligned classes. The converter instantiated the “System–Equipment” hierarchy and automatically established spatial and functional relationships using the “brick:hasLocation” and “brick:feeds” classes, respectively.

Document nodes were introduced for each of the four key files associated with every building: the IFC model, the IEDB, the EBDB, and the MEP inventory. Each document was uniquely identified by its file path and linked to both its corresponding building and time series database. Finally, two btwin:KPISet nodes were generated for each building via custom Python routines that queried the KG and the time-series repositories to calculate the metrics.

After all entities were integrated, the KG of the main building contained one site, one building, three storeys, 31 spaces, four sensors, two meters, eight sensor points, four meter points, and four document nodes. The MEP layer comprises three systems (HVAC and DHW) and 39 equipment entities, including radiators, fan-coil units, and boilers. A view of the resulting KG for the main building is presented in Figure 7.

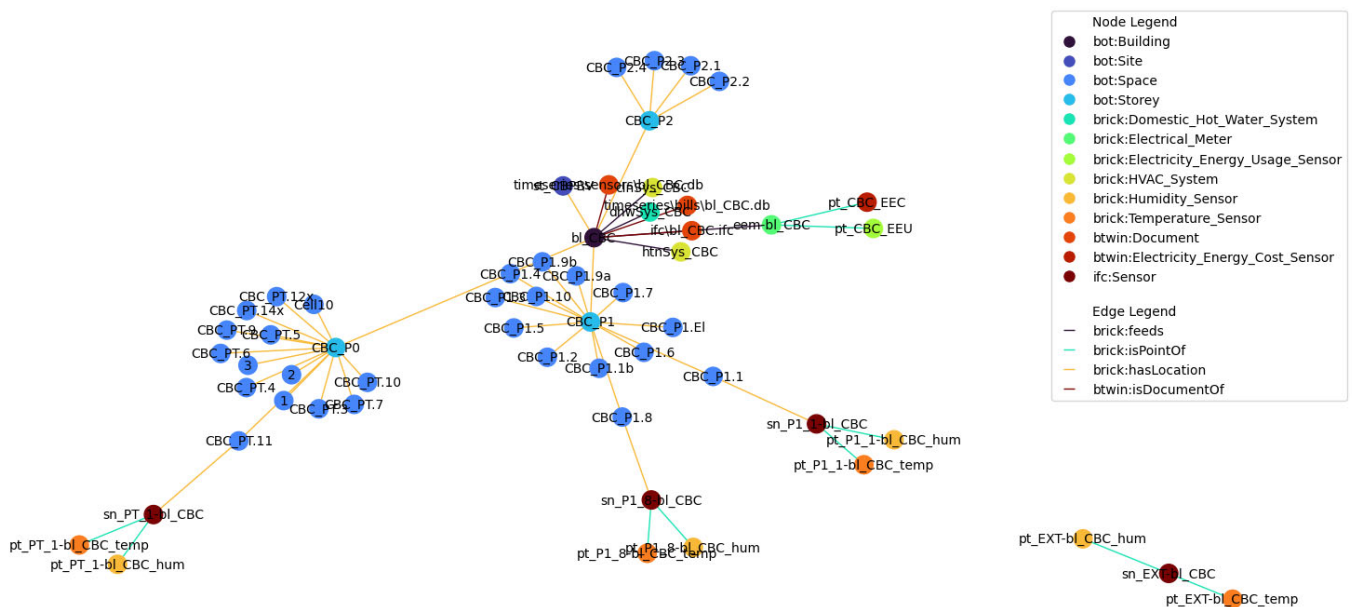


Figure 7. KG representing the main building of the Capo di Bove complex.

4.6. Development of a Web Application

After the datasets were linked in the KG, two interactive dashboards were created using Dash and Plotly: one focused on energy consumption and the other on indoor microclimatic conditions (Figure 8).

Each dashboard integrates four tightly coupled modules. To validate this structure, we conducted a series of participatory meetings with the facility managers, gathering feedback on the visualisations and integrating their requests for greater clarity of the graphical user interface.

- A filter panel queries the graph in real time to populate drop-down menus with available sites, buildings, and, in the environmental dashboard, individual spaces; it also provides a time-range picker that restricts the analysis to user-defined periods. Whenever a filter changes, a query fetches the relevant nodes from the graph and their associated time-series records to provide the correct data for display.
- An embedded IFC viewer offers both plan and three-dimensional representation of the building, allowing users to highlight spaces that satisfy the current filters and thereby link numerical indicators to their spatial context.
- A scorecard summarises key performance indicators, such as cumulative electricity use or mean indoor temperature, calculated on demand by Dash callbacks. Each callback extracts the filtered time-series data, applies the selected aggregation, and refreshes the display.
- Finally, an interactive time-series plot presents the underlying sensor or meter data. The plot updates automatically in response to filter adjustments and supports aggregation at yearly, monthly, daily, or hourly resolutions, with statistical operators including minimum, maximum, mean, and sum.

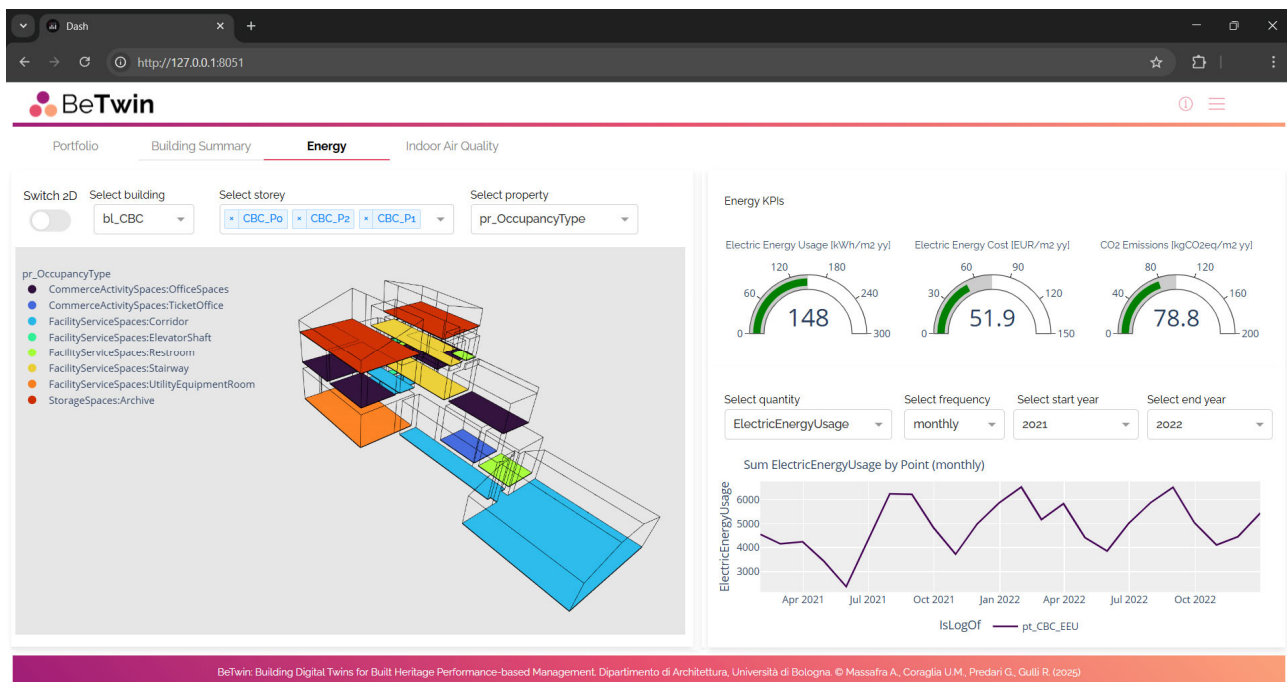


Figure 8. Prototype dashboard to access and visualise performance data.

5. Discussion

5.1. Paper Contribution

This study proposes an FDM methodology for systematically integrating building-performance information coming from heterogeneous sources into a federated KG. This KG serves as an intermediate, semantic layer between the raw data collected during the data collection campaign and a web dashboard application, providing easy access to building performance data and thereby forming a prototype DSS for energy management.

The paper's contribution is articulated below according to five key perspectives.

Data integration: The presented workflow enabled the integration of data from heterogeneous sources, addressing the fragmentation of data. In particular, it overcame the challenge of integrating dynamic data, such as energy consumption and environmental monitoring, with HBIM models, which are inherently static. This setup aligns the developed system with a DT perspective, establishing the foundational data management work necessary for building such a system when the facility is equipped with a sensing network capable of enabling continuous data flow.

Reproducibility. The data modelling and integration phases were developed according to an FOF, which was iteratively systematised and validated thanks to this and previous studies. Following the FOF, the methodology can be easily replicated across different buildings of the asset, ensuring information consistency at the portfolio level.

Scalability. The approach can be scaled to include additional input datasets by simply creating Python connectors that map the native structures of these databases to JSON-LD graph notation. In this view, the methodology has two main advantages. First, it favours open-source formats and tools such as IfcOpenShell for IFC connectors, but also Eppy for EnergyPlus, and Topologicpy for Topologic modelling, as documented in other works [37]. Second, data remains in its native environment until it is required for calculations. This avoids information redundancy from a computational standpoint while ensuring interoperability among various models/databases, each of which performs only the specific tasks for which it was designed.

Extensibility. Data management through graph structures enables an extensible, incremental approach. This incrementality can occur in two ways. First, additional

subdomains can be incorporated into the methodology by expanding the proposed FOF with ontologies from other relevant domains (e.g., fire safety, structural performance, air quality)—thereby avoiding the creation of entirely new ontological frameworks from scratch and encouraging alignment with existing ones. Second, new nodes can be added to the KG in subsequent phases, connected to existing ones, as the asset’s digitalisation progresses or new information needs emerge during the lifecycle management of the asset.

Accessibility. Data is visualised through a user-friendly interface, allowing users without IT skills to interact with the most relevant information. Moreover, graph-based data management offers significant potential for semantic integration with AI systems, such as Large Language Models [42]. This synergy could make it feasible to develop connectors with LLMs, allowing information to be retrieved via natural language queries and subsequently become more accessible.

5.2. Limitations and Future Developments

Despite the advantages mentioned, several critical issues have emerged that will guide future developments of the research.

First, even if made with low LOD, BIM modelling requires significant resources. If the methodology is to be applied to an entire building portfolio, it could be desirable to streamline and standardise it through automated or semi-automated methods [43,44].

Second, it is necessary to develop more representative KPIs that enable energy-related comparisons across different buildings. These KPIs should consider factors such as building age, functional and occupancy characteristics, which vary from one building to another in PAAA. More extensive IEQ data and additional information are also needed to perform thermal comfort assessments (e.g., PMV, PPD), covering at least 1 year of monitoring.

Third, currently, dynamic data is limited to a specific time window. The efficiency of SQL-based solutions remains to be tested as data volume increases. Specialised time-series databases—such as InfluxDB—may become necessary, requiring the development of new connectors to scale up the solution effectively.

Finally, to achieve a complete DT, both hardware and software architectures must be scaled. This includes deploying more robust databases using cloud services and equipping buildings with larger sensor networks, which must also be non-invasive given their use in cultural heritage contexts.

6. Conclusions

This research is positioned at the intersection of the ecological and digital transitions in the built environment, with a particular focus on historic assets. Specifically, it addresses the data integration gap in the AECO sector within the processes of structuring performance data required for energy audits of existing buildings. This challenge is especially evident in the case of heritage assets, where large volumes of information need to be collected from diverse sources and stakeholders, each using different formats (including unstructured sources).

The study proposes an FDM approach to link unconnected information and systematically transform it into usable knowledge.

The proposed methodology, focusing on the topic of building performance, was demonstrated through a case study located in Italy, representative of a broader sample of the national building stock. A large amount of data collected in previous works enabled the understanding of the geometrical features, spatial functions, and occupancy, as well as the typological and thermo-physical characteristics of opaque envelope components and technical systems. Additionally, indoor environmental data collected from sensors

and energy consumption records of the demonstration site were analysed. These datasets were then processed and ingested into a KG, grounded in an FOF, enabling semantic alignment across diverse information sources. Beyond serving as a semantic layer for data organisation, the KG functioned as the database underpinning a prototype decision support system (DSS). This DSS, operationalised through interactive dashboards, provided stakeholders with the capability to query KGs, thereby supporting data-driven building performance management easily.

The case study was particularly valuable for testing the methodology, drafted in previous works, demonstrating how it can bridge the gap between data availability and decision-making in a real operational environment.

Paths for further investigation were identified at the end of the paper. As future improvements, the scalability of the FDM method will be tested across a broader range of heritage typologies to ensure generalisability. The long-term governance of semantically enriched datasets, including updating protocols, data ownership, and institutional integration, will also be explored. Furthermore, research activities will be coupled with tasks aimed at quantifying the economic, social, and cultural benefits associated with digital decision environments, which are essential for informing investment decisions and ensuring the sustainability of digital innovations in the sector.

Future technological developments will also envisage evolving the system towards a fully operational DT, enabling bidirectional data exchange between physical and virtual assets, and integration with AI to enhance predictive capabilities and facilitate user-friendly, customised access to building data.

Author Contributions: Conceptualization, A.M. and U.M.C.; methodology, A.M. and U.M.C.; software, A.M. and U.M.C.; validation, S.D.T. and D.P.; formal analysis, S.D.T. and D.P.; investigation, S.D.T. and D.P.; data curation, A.M., U.M.C., S.D.T. and D.P.; writing—original draft preparation, A.M., U.M.C., S.D.T. and D.P.; writing—review and editing, A.M., U.M.C., S.D.T. and D.P.; visualisation, A.M., U.M.C., S.D.T. and D.P.; supervision, A.M. and U.M.C.; project administration, A.M. and U.M.C.; funding acquisition, A.M. All authors have read and agreed to the published version of the manuscript.

Funding: The methodological development of this research was funded by the BeTwin project (“Building Digital Twins for Built Heritage Performance-Based Management”) under the Alma-Value 2023 grant of the University of Bologna (funded by the NextGenerationEU programme). Data collection, case study investigation, and energy audits/analyses were founded in the Program Agreement between the Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA) and the Ministry of Economic Development (now Ministry of Ecological Transition) for Electric System Research, in the framework of its Implementation Plan for 2022–2024. In particular, the activity is included in the Project 1.5 “Highly efficient buildings for the energy transition”. Project administration: Giovanni Puglisi.

Data Availability Statement: The data used in this study are not publicly available as they may contain sensitive information belonging to the case study’s owner, i.e., Ente Parco Archeologico dell’Appia Antica. However, interested parties may contact the corresponding author to obtain the templates used to support the data integration method, which contain only generic and fictitious data.

Acknowledgments: The authors thank Ente Parco Archeologico dell’Appia Antica in the figures of Simone Quilici, Michele Reginaldi and Fabiana Volpe for their support and for the activities inherent to the execution of the scientific collaboration between the Department of Architecture of the University of Bologna, ENEA and the park.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AECO	Architecture, Engineering, Construction, and Operations
AHU	Air Handling Unit
AI	Artificial Intelligence
BIM	Building Information Modelling
BOT	Building Topology Ontology
DT	Digital Twin
EBDB	Energy Bill Database
EC	Energy Cost
EM-KPI	Energy Management Key Performance Indicator Ontology
EU	Energy Usage
FDM	Federated Data Modelling
FOF	Federated Ontology Framework
FM	Facility Management
GIS	Geographic Information System
HVAC	Heating, Ventilation and Air Conditioning
HBIM	Heritage Building Information Modelling
IEDB	Indoor Environment Database
IFC	Industry Foundation Classes
IoT	Internet of Things
KG	Knowledge Graph
LD	Linked Data
LLM	Large Language Model
LOD	Level of Development
LPG	Labelled Property Graph
MEP	Mechanical, Electrical, Plumbing
PAAA	Parco Archeologico dell'Appia Antica
RDF	Resource Description Framework
SOSA	Sensor, Observation, Sample, and Actuator
SSN	Semantic Sensor Network Ontology
UID	Unique Identifier

Appendix A

Appendix A.1. Building Envelope

The opaque vertical components of the case study consist of mixed masonry walls built using tuff, flint, brick, and reused Roman and medieval elements. On the southeast façades, remnants of the Roman cistern's original flint rubble masonry are still visible, overlaid by early 20th-century structural additions. Wall thicknesses vary between approximately 75 cm (at the cistern walls) and 40 cm in other parts of the building.

The roofs are predominantly pitched, with wood and brick elements covered with traditional clay tiles. Suspended ceilings have also been installed under the roof slopes, where space is used for the location of the HVAC systems. A small flat walkable roof, presumably composed of hollow brick and concrete, covers the central corridor and lacks thermal insulation.

Due to the absence of archival documentation, two typologies are hypothesised for the floor slabs: one supported directly by the ancient cistern's masonry and another consisting of a concrete slab laid over a ventilated crawl space filled with stone. Interior floor finishes are primarily in terracotta tiles.

Overall, the building envelope reflects the layering of historical construction phases, with heterogeneous materials and techniques. No thermal insulation has been identified in the envelope components, resulting in varying and generally poor thermal performance.

Appendix A.2. Mechanical and Electrical Systems

The main building is equipped with an autonomous HVAC system based on a reversible hydronic heat pump, located outdoors and directly connected to the distribution network via two single-stage twin-circulation pumps in the ground-floor mechanical room. These pumps supply the thermal carrier fluid to both fan coil units and two air handling units (AHUs), one placed in the mechanical room and one in the attic above the ground floor.

A total of 19 wall-mounted fan coil units were identified during the on-site survey, located in corridors and offices, except for the conference and exhibition halls, which are conditioned via air ducts connected to the two AHUs. The thermal output of the fan coils ranges from 3.11 kW to 6.53 kW (at 40–45 °C inlet water temperature).

The AHU in the mechanical room serves the exhibition space, while the unit in the suspended ceiling, accessible via a hatch in the ground-floor bathroom, serves the conference room. In both cases, supply diffusers are ceiling-mounted and concealed within plasterboard coves, while return grilles are wall-mounted near floor level (visible only in the conference hall). Each AHU manages an air flow rate of 2000 m³/h.

Additionally, a portable air conditioning unit is installed in the first-floor server room to provide local cooling due to high internal heat loads.

Three direct-expansion reversible heat pumps condition the outbuilding. The outdoor units are installed behind the building.

The domestic hot water system includes seven electric boilers, five in the main building and two in the outbuilding.

Interior lighting in the main building primarily consists of fixtures with one or two fluorescent tubes (18 or 36 W), especially in corridors and shared areas. LED lighting was observed in select areas (e.g., the annex reception, bathrooms, and some offices), while halogen lamps, primarily as desk lamps, were in smaller quantities.

Appendix A.3. Energy Bills

The Capo di Bove complex operates using electricity as its sole energy carrier for all end uses, including space heating, cooling, domestic hot water, and lighting. Although the same utility provider supplies both buildings, they are served through two distinct electrical contracts. Electricity consumption data were obtained from utility bills made available by the site Administration for the 2021–2022 years. In 2021, the highest energy consumption in the main building—approximately 6200 kWh—was recorded in July and August. In 2022, peak values—slightly exceeding 6500 kWh—were observed in January and August (Figure A1). Regarding the outbuilding (Figure A2), in 2022, January recorded the highest consumption, whereas in 2021 it was approximately 15% lower than the peak value, which occurred in December. The months with the lowest consumption were May 2021 and October 2022.

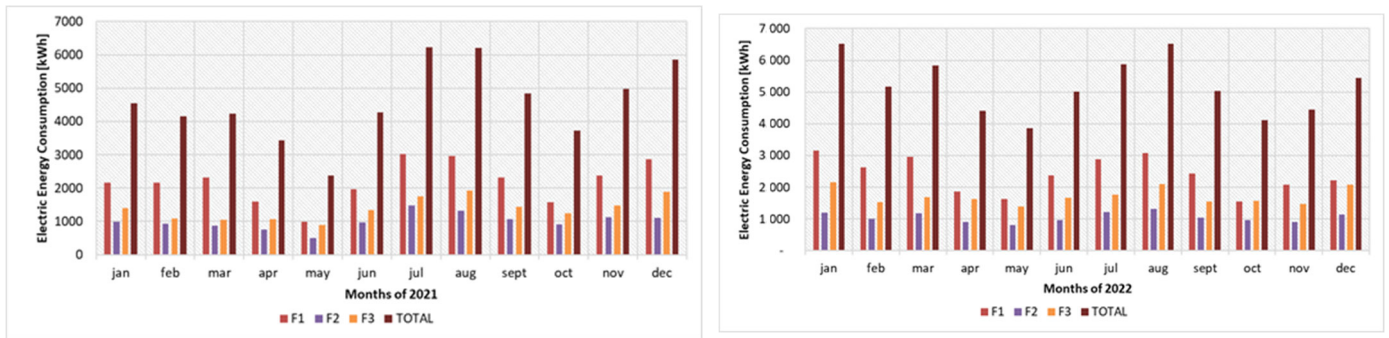


Figure A1. Electric energy consumption data from utility bills for the main building of Capo di Bove, categorised by time-of-use tariff bands (F1, F2, F3, total).

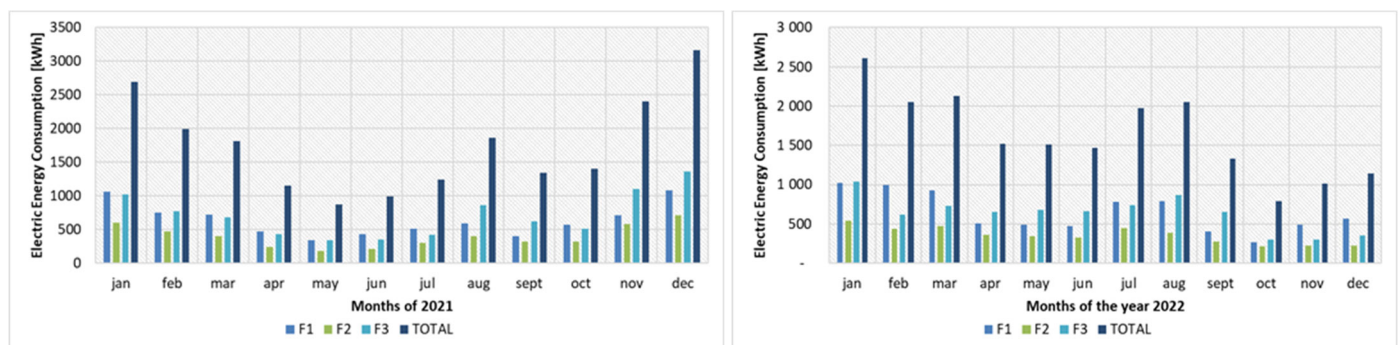


Figure A2. Electric energy consumption data from utility bills for the outbuilding, categorised by time-of-use tariff bands (F1, F2, F3, total).

Appendix A.4. Microclimatic Data

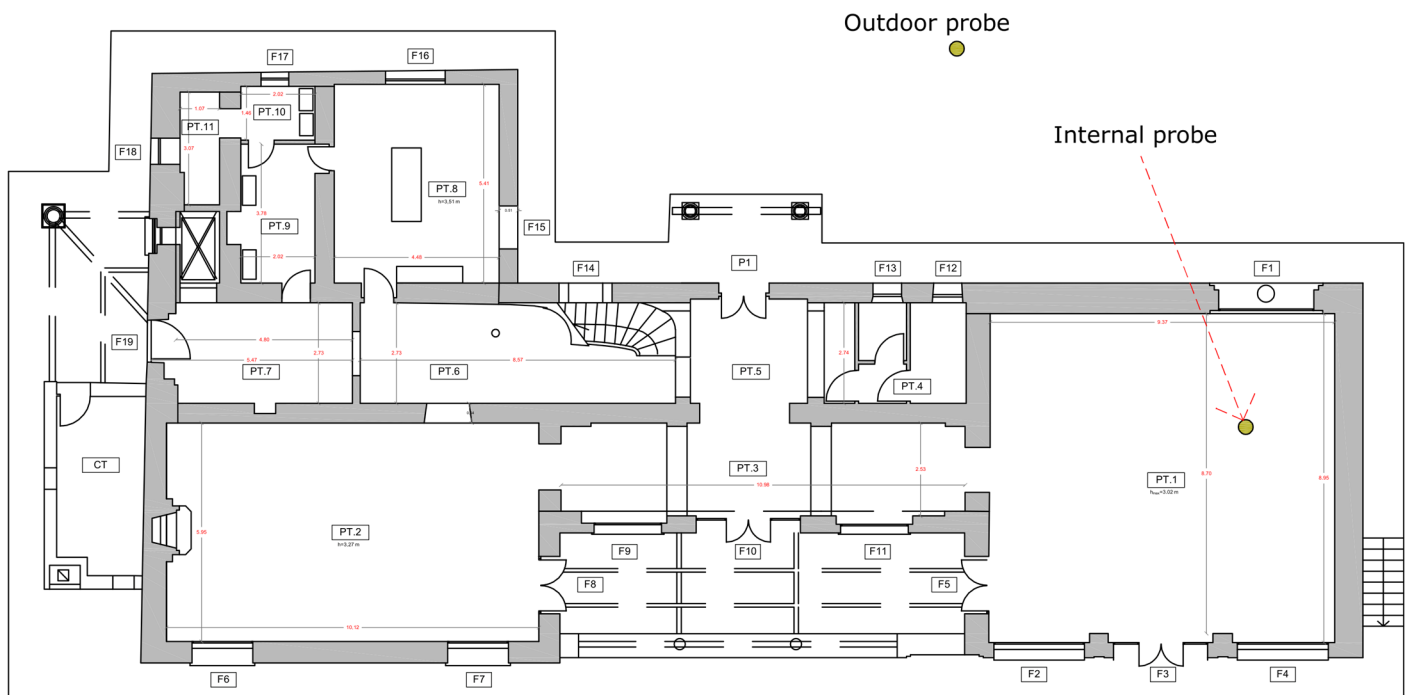
The building was also subjected to environmental monitoring aimed at assessing the actual indoor microclimatic conditions in relation to outdoor conditions, both with HVAC systems turned off and on. This type of investigation enabled a more detailed analysis of the building's actual energy behaviour under varying external climatic conditions, highlighting potential issues with the air-conditioning system's operation or the need for climate control in its absence.

The measurement equipment consisted of data loggers, which allow for the monitoring of temperature and relative humidity in both indoor and outdoor environments. For indoor monitoring, devices capable of measuring temperatures ranging from $-25\text{ }^{\circ}\text{C}$ to $+85\text{ }^{\circ}\text{C}$ and relative humidity from 0% to 95% were employed. For outdoor monitoring, the used device was capable of recording air temperature within the same range ($-25\text{ }^{\circ}\text{C}$ to $+85\text{ }^{\circ}\text{C}$) and relative humidity from 0% to 100%.

Both instruments can record minimum, average, and maximum values of temperature and humidity, with sampling intervals ranging from 1 s to 10 days, and offer an accuracy of $\pm 0.4\text{ }^{\circ}\text{C}$ and $\pm 3.0\%$ (values referred to a temperature of $25\text{ }^{\circ}\text{C}$).

An example of the positioning of the measuring instruments and the monitored rooms is shown in Figure A3, while the results of the related monitoring campaign—conducted over approximately 38 days between 24 May and 1 July 2023—are presented in Figure A4 (indoor air temperature) and Figure A5 (relative humidity).

Ground Floor



First Floor

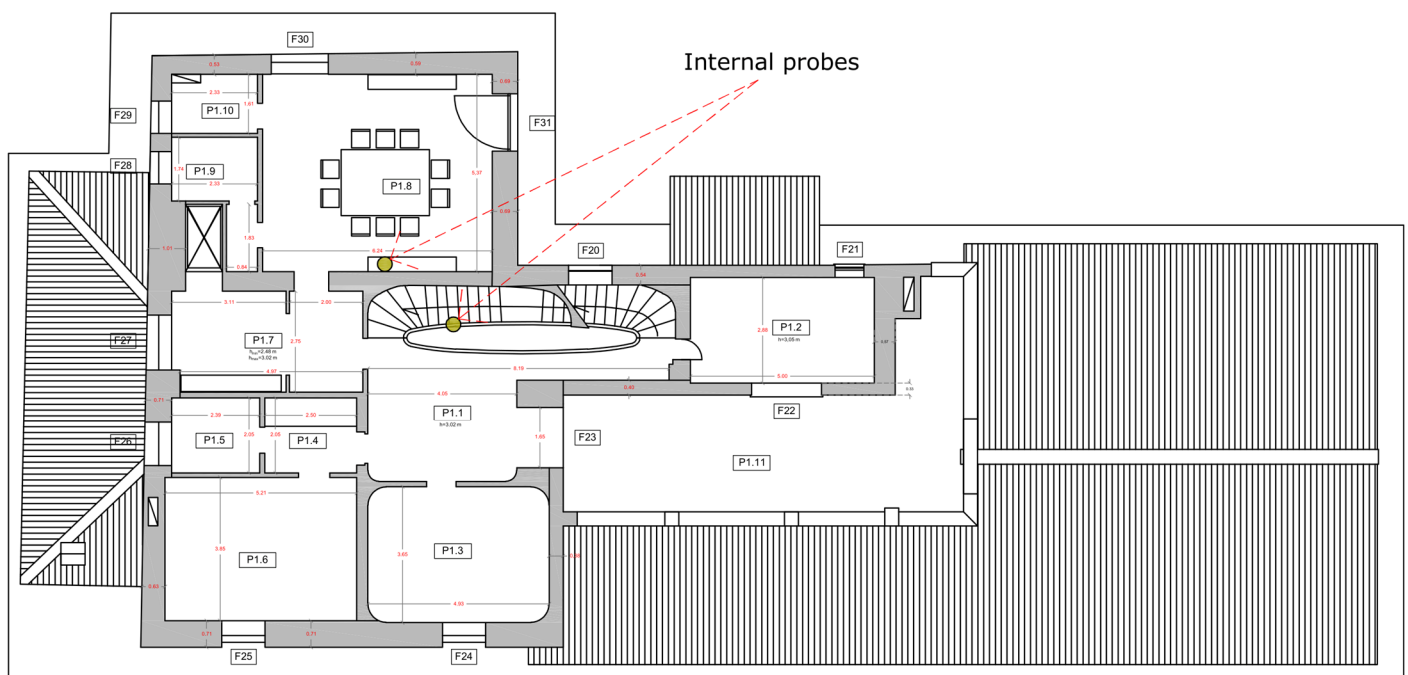


Figure A3. Spaces under investigation and placement of monitoring equipment for environmental data collection at main building (ground floor on the left and first floor on the right).

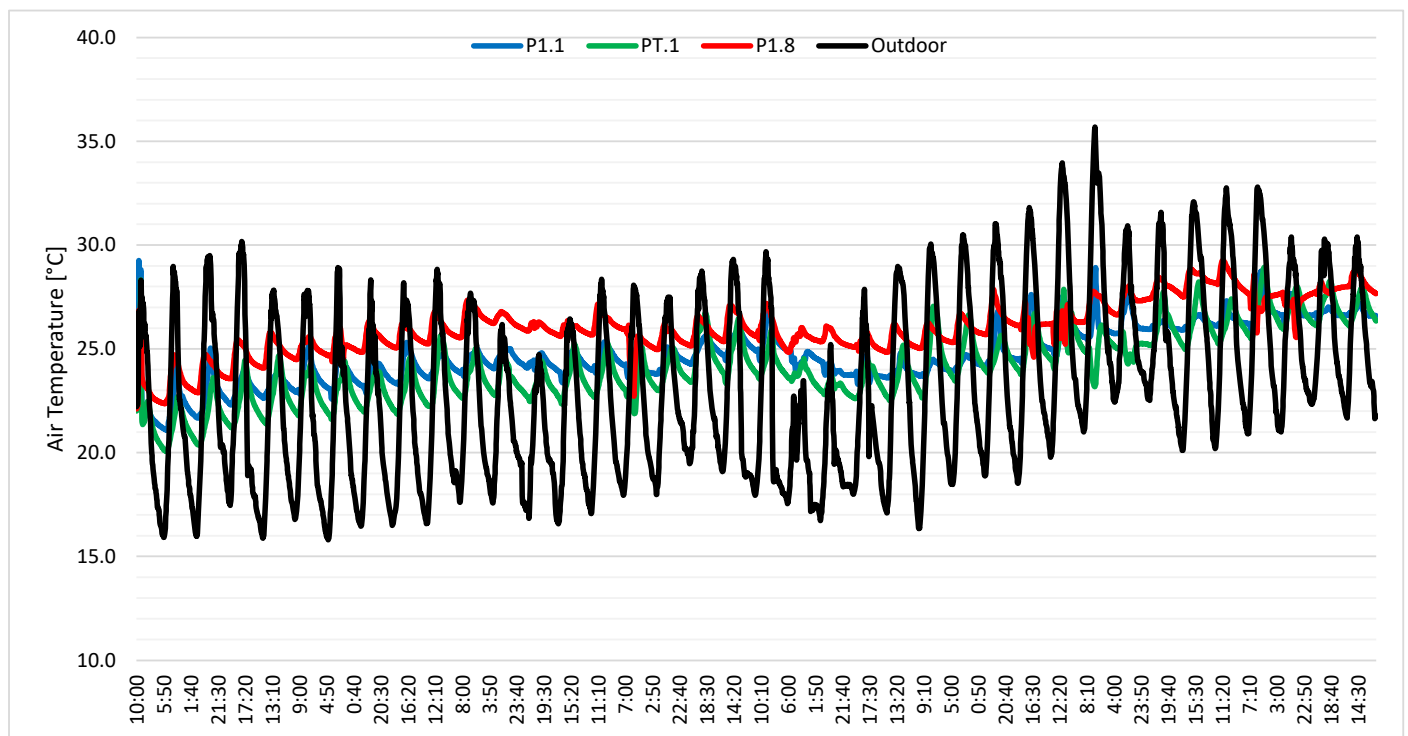


Figure A4. Data collected during the experimental environmental monitoring campaign at Capo di Bove: indoor and outdoor temperature.

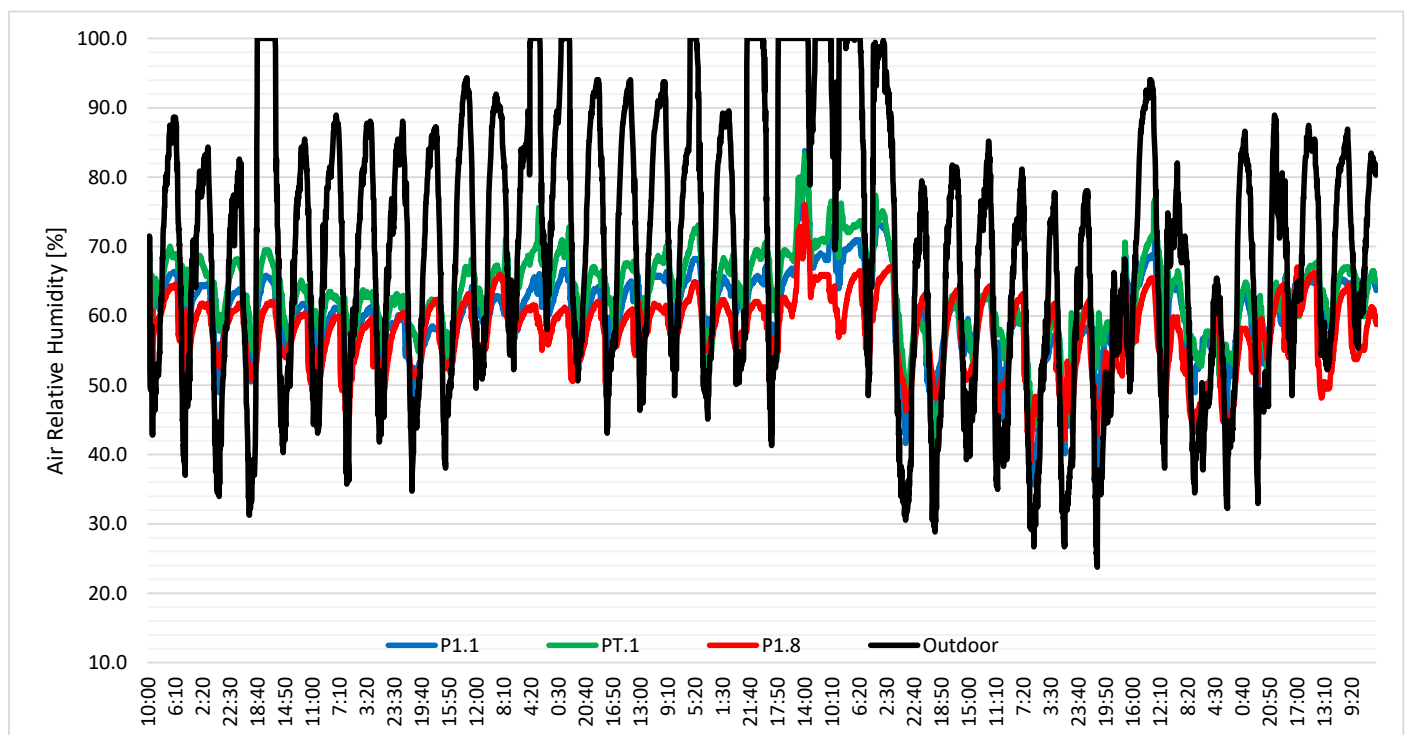


Figure A5. Data collected during the experimental environmental monitoring campaign at Capo di Bove: indoor and outdoor relative humidity.

The indoor temperature recorded in the conference centre, the staircase, and the largest office displays a rising trend during the initial monitoring period, despite relatively stable outdoor temperatures. This behaviour is likely due to internal heat gains and the HVAC system being switched off. In contrast, during the mid and final phases of the monitoring period (June), an increase in outdoor temperature corresponds with a rise in indoor

climatic conditions, although less pronounced, thanks to the activation of the air-conditioning system during that time.

Interestingly, the space with the most critical microclimatic conditions is one of the offices on first floor (average temperature on working days: $26\text{ °C} \pm 1.3\text{ °C}$; average relative humidity: $57.5\% \pm 19.2\%$), which faces northeast. This suggests that the most significant thermal discomfort may be associated with a higher occupancy rate in that specific room. The staircase, despite being a central space, shows intermediate microclimatic conditions compared to the other monitored areas.

References

1. Al-Addous, M.; Albatayneh, A. Knowledge gap with the existing building energy assessment systems. *Energy Explor. Exploit.* **2020**, *38*, 783–794. <https://doi.org/10.1177/0144598719888100>.
2. Takva, Ç.; Takva, F.G.; Çakıcı, F.Z. Thermal Analysis of the Building Envelope with Infrared Thermography and Simulation in Educational Buildings in the Cold Climate Region. *Buildings* **2025**, *15*, 1759. <https://doi.org/10.3390/buildings15111759>.
3. Patel, D.; Estevam Schmiedt, J.; Röger, M.; Hoffschmidt, B. A Model Calibration Approach to U-Value Measurements with Thermography. *Buildings* **2023**, *13*, 2253. <https://doi.org/10.3390/buildings13092253>.
4. Eun Jung, D.; Yoo, S.; Ho Lee, K.; Kim, J. In-situ measurement of thermal transmittance of building walls: Evaluation of stored heat flux in heavy-weight walls during the cooling season. *Energy Build.* **2024**, *325*, 114981. <https://doi.org/10.1016/j.enbuild.2024.114981>.
5. Tardy, F. Methodology for estimating building thermal resistance and heat capacity values in-situ using exterior measurements and meteorological data. *J. Build. Eng.* **2025**, *101*, 111768. <https://doi.org/10.1016/j.job.2025.111768>.
6. Merino, J.; Xie, X.; Moretti, N.; Chang, J.Y.; Parlikad, A.K. Data integration for digital twins in the built environment. In Proceedings of the 2022 European Conference on Computing in Construction, Rhodes, Greece, 24–26 July 2022.
7. Abuimara, T.; Hobson, B.W.; Gunay, B.; O'Brien, W.; Kane, M. Current state and future challenges in building management: Practitioner interviews and a literature review. *J. Build. Eng.* **2021**, *41*, 102803. <https://doi.org/10.1016/j.job.2021.102803>.
8. Petri, I.; Rezgui, Y.; Ghoroghi, A.; Alzahrani, A. Digital twins for performance management in the built environment. *J. Ind. Inf. Integr.* **2023**, *33*, 100445. <https://doi.org/10.1016/j.jii.2023.100445>.
9. Moretti, N.; Xie, X.; Merino Garcia, J.; Chang, J.; Kumar Parlikad, A. Federated Data Modeling for Built Environment Digital Twins. *J. Comput. Civ. Eng.* **2023**, *37*, 04023013. <https://doi.org/10.1061/JCCEE5.CPENG-4859>.
10. Massafra, A.; Coraglia, U.M.; Predari, G.; Gulli, R. Graph-Based Digital Decision Support Systems: Introducing BTwin, a Toolkit for Building Performance Management. In Proceedings of the 2024 European Conference on Computing in Construction, Chania, Crete, Greece, 15–17 July 2024.
11. Di Turi, S.; Palladino, D.; Caffari, F.; Centi, G.; Margiotta, F.; Ronchetti, L.; Signoretti, P.; Volpe, L. *Energy Efficiency of Complexes of Historic and Protected Buildings: Characterisation of Real Case Studies (LA1.4)*; ENEA—Italian National Agency for New Technologies, Energy and Sustainable Economic Development: Rome, Italy, 2023.
12. Yu, Y.; Raed, A.A.; Peng, Y.; Pottgiesser, U.; Verbree, E.; Van Oosterom, P. How digital technologies have been applied for architectural heritage risk management: A systemic literature review from 2014 to 2024. *NPJ Herit. Sci.* **2025**, *13*, 45. <https://doi.org/10.1038/s40494-025-01558-5>.
13. Murphy, M.; McGovern, E.; Pavia, S. Historic building information modelling (HBIM). *Struct. Surv.* **2009**, *27*, 311–327. <https://doi.org/10.1108/02630800910985108>.
14. Penjor, T.; Banhashemi, S.; Hajirasouli, A.; Golzad, H. Heritage building information modeling (HBIM) for heritage conservation: Framework of challenges, gaps, and existing limitations of HBIM. *Digit. Appl. Archaeol. Cult. Herit.* **2024**, *35*, e00366. <https://doi.org/10.1016/j.daach.2024.e00366>.
15. Lovell, L.J.; Davies, R.J.; Hunt, D.V.L. The Application of Historic Building Information Modelling (HBIM) to Cultural Heritage: A Review. *Heritage* **2023**, *6*, 6691–6717. <https://doi.org/10.3390/heritage6100350>.
16. Puerto, A.; Castañeda, K.; Sánchez, O.; Peña, C.A.; Gutiérrez, L.; Sáenz, P. Building information modeling and complementary technologies in heritage buildings: A bibliometric analysis. *Results Eng.* **2024**, *22*, 102192. <https://doi.org/10.1016/j.rineng.2024.102192>.
17. Hou, H.; Lai, J.H.K.; Wu, H.; Wang, T. Digital twin application in heritage facilities management: Systematic literature review and future development directions. *Eng. Constr. Archit. Manag.* **2024**, *31*, 3193–3221. <https://doi.org/10.1108/ECAM-06-2022-0596>.

18. Pauwels, P.; Zhang, S.; Lee, Y.-C. Semantic web technologies in AEC industry: A literature overview. *Autom. Constr.* **2017**, *73*, 145–165. <https://doi.org/10.1016/j.autcon.2016.10.003>.
19. Curry, E.; O'Donnell, J.; Corry, E.; Hasan, S.; Keane, M.; O'Riain, S. Linking building data in the cloud: Integrating cross-domain building data using linked data. *Adv. Eng. Inform.* **2013**, *27*, 206–219. <https://doi.org/10.1016/j.aei.2012.10.003>.
20. Teclaw, W.; O'Donnell, J.; Kukkonen, V.; Pauwels, P.; Labonnote, N.; Hjelseth, E. Federating cross-domain BIM-based knowledge graph. *Adv. Eng. Inform.* **2024**, *62*, 102770. <https://doi.org/10.1016/j.aei.2024.102770>.
21. Ramonell, C.; Chacón, R.; Posada, H. Knowledge graph-based data integration system for digital twins of built assets. *Autom. Constr.* **2023**, *156*, 105109. <https://doi.org/10.1016/j.autcon.2023.105109>.
22. Paulheim, H. Knowledge graph refinement: A survey of approaches and evaluation methods. *Semant. Web* **2016**, *8*, 489–508. <https://doi.org/10.3233/SW-160218>.
23. Wang, Z.; Ying, H.; Sacks, R.; Borrmann, A. CBIM: A Graph-based Approach to Enhance Interoperability Using Semantic Enrichment. *arXiv* **2023**. <https://doi.org/10.48550/ARXIV.2304.11672>.
24. Gruber, T.R. A translation approach to portable ontology specifications. *Knowl. Acquis.* **1993**, *5*, 199–220. <https://doi.org/10.1006/knac.1993.1008>.
25. EC3 M&S Committee Built Environment Ontology Lookup Service. 2025. Available online: https://ec-3.org/publications/conference/paper/?id=EC32025_353 (accessed on 28 September 2025).
26. buildingSMART International IFC 4.3.2.20251007 (IFC4X3_ADD2) 2025. Available online: <https://ifc43-docs.standards.buildingsmart.org/> (accessed on 28 September 2025).
27. Rasmussen, M.H.; Pauwels, P.; Hviid, C.A.; Karlshøj, J. Proposing a Central AEC Ontology That Allows for Domain Specific Extensions. In Proceedings of the Lean and Computing in Construction Congress-Volume 1: Proceedings of the Joint Conference on Computing in Construction, Heraklion, Crete, Greece, 9–12 July 2017; Heriot-Watt University: Heraklion, Crete, Greece, 2017; pp. 237–244.
28. Balaji, B.; Bhattacharya, A.; Fierro, G.; Gao, J.; Gluck, J.; Hong, D.; Johansen, A.; Koh, J.; Ploennigs, J.; Agarwal, Y.; et al. Brick : Metadata schema for portable smart building applications. *Appl. Energy* **2018**, *226*, 1273–1292. <https://doi.org/10.1016/j.apenergy.2018.02.091>.
29. Compton, M.; Barnaghi, P.; Bermudez, L.; García-Castro, R.; Corcho, O.; Cox, S.; Graybeal, J.; Hauswirth, M.; Henson, C.; Herzog, A.; et al. The SSN ontology of the W3C semantic sensor network incubator group. *J. Web Semant.* **2012**, *17*, 25–32. <https://doi.org/10.1016/j.websem.2012.05.003>.
30. Janowicz, K.; Haller, A.; Cox, S.J.D.; Phuoc, D.L.; Lefrancois, M. SOSA: A Lightweight Ontology for Sensors, Observations, Samples, and Actuators. *arXiv* **2018**. <https://doi.org/10.48550/ARXIV.1805.09979>.
31. Li, Y.; García-Castro, R.; Mihindukulasooriya, N.; O'Donnell, J.; Vega-Sánchez, S. Enhancing energy management at district and building levels via an EM-KPI ontology. *Autom. Constr.* **2019**, *99*, 152–167. <https://doi.org/10.1016/j.autcon.2018.12.010>.
32. Chamari, L.; Petrova, E.; Pauwels, P. A web-based approach to BMS, BIM and IoT integration: A case study. In Proceedings of the CLIMA 2022 Conference 2022 The 14th REHVA HVAC World Congress, Rotterdam, The Netherlands, 22–25 May 2022. <https://doi.org/10.34641/clima.2022.228>.
33. Niccolucci, F.; Felicetti, A. Digital Twin Sensors in Cultural Heritage Ontology Applications. *Sensors* **2024**, *24*, 3978. <https://doi.org/10.3390/s24123978>.
34. Hosamo, H.H.; Nielsen, H.K.; Kraniotis, D.; Svennevig, P.R.; Svidt, K. Digital Twin framework for automated fault source detection and prediction for comfort performance evaluation of existing non-residential Norwegian buildings. *Energy Build.* **2023**, *281*, 112732. <https://doi.org/10.1016/j.enbuild.2022.112732>.
35. Falcone, M.; Origlia, A.; Campi, M.; Di Martino, S. From Architectural Survey to Continuous Monitoring: Graph-Based Data Management for Cultural Heritage Conservation with Digital Twins. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2021**, *XLIII-B4-2021*, 47–53. <https://doi.org/10.5194/isprs-archives-XLIII-B4-2021-47-2021>.
36. Massafra, A.; Costantino, C.; Predari, G.; Gulli, R. Building Information Modeling and Building Performance Simulation-Based Decision Support Systems for Improved Built Heritage Operation. *Sustainability* **2023**, *15*, 11240. <https://doi.org/10.3390/su151411240>.
37. Massafra, A.; Coraglia, U.M.; Predari, G.; Gulli, R. Digital Decision Support System Prototyping for Building Performance Analysis and Management. In *Proceedings of the 11th International Conference of Società Scientifica di Architettura Tecnica (Scientific Society of Architectural Engineering), Palermo, Italy, 12–15 June 2024*; Corrao, R., Campisi, T., Colajanni, S., Saeli, M., Vinci, C., Eds.; Springer Nature Switzerland: Cham, Switzerland, 2025; Volume 611, pp. 489–506, ISBN 978-3-031-71862-5.

38. Di Turi, S.; Palladino, D.; Murano, G.; Centi, G.; Ronchetti, L.; Signoretti, P. *Energy Efficiency of Complexes of Historic and Protected Buildings: Energy Audit of Case Studies and Assessment of the Retrofit Interventions (LA1.5)*; ENEA–Italian National Agency for New Technologies, Energy and Sustainable Economic Development: Rome, Italy, 2024.
39. Pritoni, M.; Paine, D.; Fierro, G.; Mosiman, C.; Poplawski, M.; Saha, A.; Bender, J.; Granderson, J. Metadata Schemas and Ontologies for Building Energy Applications: A Critical Review and Use Case Analysis. *Energies* **2021**, *14*, 2024. <https://doi.org/10.3390/en14072024>.
40. Jabi, W.; Chatzivasileiadi, A. Topologic: Exploring Spatial Reasoning Through Geometry, Topology, and Semantics. In *Advances in Science, Technology & Innovation*; Springer International Publishing: Cham, Switzerland, 2021.
41. CSI. *OmniClass: A Strategy for Classifying the Built Environment. Introduction and User's Guide. Edition 2.1. 22*; CIS: Paducah, KY, USA, 2019.
42. Arazzi, M.; Ligari, D.; Nicolazzo, S.; Nocera, A. Augmented Knowledge Graph Querying leveraging LLMs. *arXiv* **2025**. <https://doi.org/10.48550/ARXIV.2502.01298>.
43. Massafra, A.; Jabi, W.; Gulli, R. Topological BIM for building performance management. *Autom. Constr.* **2024**, *166*, 105628. <https://doi.org/10.1016/j.autcon.2024.105628>.
44. Bortoluzzi, B.; Efremov, I.; Medina, C.; Sobieraj, D.; McArthur, J.J. Automating the creation of building information models for existing buildings. *Autom. Constr.* **2019**, *105*, 102838. <https://doi.org/10.1016/j.autcon.2019.102838>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.