Optimizing Carbon Footprint & Energy Performance for the Sustainability of Historic Buildings using Knowledge Graphs & Digital Twins

Alexandros Vassiliades[1,*,†], Anastasios I. Karageorgiadis[1,†], Elissavet Batziou[1,†], Sotiris Diplaris[1], Evangelos A. Stathopoulos[1], Nikolaos Dourvas[1], Ioannidis Konstantinos[1], Stefanos Vrochidis[1] and Ioannis Kompatsiaris[1]

[1]*Information Technologies Institute, Center for Research and Technology Hellas, Thessaloniki, Greece*

Abstract

Cultural heritage preservation is crucial for climate resilience and sustainable development, requiring innovative tools, materials, and adaptive renovation to mitigate climate risks, reduce emissions, and enhance sustainability in line with the EU Green Deal and UN Sustainable Development Goals. This paper presents ongoing work integrating Knowledge Graphs, Digital Twins, and Building Information Modeling (BIM) to optimize the carbon footprint and energy performance of historic buildings through innovative restoration materials, energy harvesting technologies, and socially-driven approaches, aligning with net-zero-carbon goals. We propose a pipeline where a Digital Twin, incorporating a BIM model, simulates a historic building's virtual representation to evaluate how different materials impact energy consumption and sustainability. The Knowledge Graph stores historical, real-time (sensor-based), and predicted weather data, enabling the Digital Twin to assess weather-driven energy performance variations and determine optimal material choices. As part of the EU-funded SINCERE project, this system provides a data-driven decision-making framework for stakeholders, supporting restoration, operation, and long-term sustainability planning for Built Cultural Heritage.

Keywords

Knowledge Graph, Digital Twin, Ontology, Information Retrieval, Weather Data, Building Information Modelling Models

1. Introduction

According to UNESCO, cultural heritage (CH) is "our legacy from the past, what we live with today, and what we pass on to future generations". Built Heritage is indisputably of unique cultural, social, environmental, and economic value. Its existence in the modern world preserves the history and culture of each nation and offers great potential to drive climate action and contribute to a climate resilient future (EU – "Green Deal"). As a matter of fact, climate change is a reality around the world and its extent and speed of change is becoming ever more evident, causing a wide range of environmental, societal, and economic impacts. Aiming to meet the goal of passing CH to the next generations, it is essential to explore and develop tools, materials, and technologies for protecting from climate change risks and enabling Built Heritage to contribute to the achievement of the Sustainable Development Goals (SDGs) as described by UN in the "2030 Agenda for Sustainable Development". Hence, renovation and rehabilitation avoid unnecessary greenhouse gas emissions, preserves the heritage capital of EU, and promotes the sustainable development of cities, including economic growth, social wellbeing, and environmental preservation, in line with circular economy rules.

The third international Workshop on Knowledge Graphs for Sustainability - KG4S, June 2025, Portoroz, Slovenia.

*Corresponding author.
†These authors contributed equally.
valexande@iti.gr (A. Vassiliades); tassoskarag@iti.gr (A. I. Karageorgiadis); batziou.el@iti.gr (E. Batziou); diplaris@iti.gr (S. Diplaris); estathop@iti.gr (E. A. Stathopoulos); ndourvas@iti.gr (N. Dourvas); kioannid@iti.gr (I. Konstantinos); stefanos@iti.gr (S. Vrochidis); ikom@iti.gr (I. Kompatsiaris)
0000-0003-4569-503X (A. Vassiliades); 0009-0006-0878-5131 (A. I. Karageorgiadis); 0000-0003-2645-939X (E. Batziou); 0000-0002-9969-6436 (S. Diplaris); 0000-0003-3713-5833 (E. A. Stathopoulos); 0000-0002-6202-4556 (N. Dourvas); 0000–0001–6767–8762 (I. Konstantinos); 0000-0002-2505-9178 (S. Vrochidis); 0000-0001-6447-9020 (I. Kompatsiaris)

In this paper we present, a system that combines Knowledge Graphs (KGs), Digital Twins (DTs), and Building Information Modelling (BIM) to improve the sustainability and energy efficiency of historic buildings. By integrating advanced restoration materials, energy harvesting technologies, and socially-driven strategies, our approach supports the goal of net-zero-carbon buildings. Our proposed pipeline uses a DT—a virtual replica of a historic building—built from BIM data. This DT allows us to test how different materials affect energy consumption and overall sustainability. A KG serves as the system's intelligent database, storing historical weather data, real-time sensor readings, and future climate predictions. By feeding this information into the DT, the system can analyse how weather conditions influence energy performance and recommend the most effective materials for renovation.

The aforementioned pipeline is part of a bigger project called SINCERE that has as main motivation to elucidate the values of Built Cultural Heritage and provide the tools for optimising the carbon footprint and energy performance of historic buildings, towards the requirements of net-zero-carbon-buildings, by utilising innovative, sustainable, and cost-effective restoration materials and practices, energy harvesting technologies, ICT tools and socially innovative approaches.

The main contribution of this paper is the pipeline itself, where the KG provides weather data along with their semantic relationships to the DT. The DT, using both the weather data and the BIM, computes the energy consumption of the building and evaluates its sustainability when various materials are applied to its surfaces (more details about the use cases will be provided in Section 3). An additional contribution is the generation of synthetic weather data for the years 1980–2020 and 2025–2065, produced using weather prediction models trained on historical weather data from locations with similar climates.

A key limitation of our approach is the dependency on prediction models for generating synthetic data. Another limitation is that our methodology has been tested only on CH buildings, meaning the pipeline might require optimizations if applied to other types of buildings, such as factories.

In Section 2 we give the related work. Next, in Section 3, we talk about the data that is mapped in the KG and the BIM files, we also discuss briefly the architecture of the KG and what is the nature of the problem that the DT has to solve, in order to achieve the optimisation in energy consumption and preservance of the building. We conclude our paper with a discussion and conclusion in Section 4.

2. Related Work

The integration of KGs, DTs, and BIM for optimizing energy efficiency in CH buildings has been an emerging area of research. Several works have explored individual aspects of these technologies; however, their combined application for sustainability and energy performance remains under explored.

A significant body of research has focused on the use of BIM for historical buildings, primarily in the context of Heritage Building Information Modeling (HBIM). HBIM provides a structured way to represent and manage heritage structures digitally [1, 2]. These models allow for efficient documentation, analysis, and conservation planning. However, HBIM alone lacks the capability to integrate dynamic environmental factors, such as weather variations and material degradation over time.

The concept of DT has gained traction in recent years, with applications spanning smart cities, industrial manufacturing, and the built environment [3, 4]. In the context of CH, DTs enable real-time monitoring and predictive analysis by integrating sensor data with virtual models [5]. While DTs have been used for predictive maintenance and structural health monitoring [6], their role in energy optimization remains an ongoing research challenge.

KGs have also been applied in various domains for semantic data integration and reasoning [7]. In the CH domain, ontologies have been developed to model heritage structures, enabling interoperability between datasets [8]. The use of KGs in sustainability-related applications is still evolving, with recent efforts focused on integrating multi-source data for decision-making [9]. However, few studies have combined KGs with DTs and BIM to create an end-to-end energy optimization pipeline.

Our work builds upon these foundations by leveraging the synergy between KGs, DTs, and BIM to enable holistic energy performance optimization for historic buildings. Unlike previous approaches that focus on static building models or limited sensor integration, our proposed pipeline dynamically incorporates historical, real-time, and predictive weather data into a DT-driven simulation framework. By integrating these elements, we aim to provide stakeholders with a comprehensive decision-support system for sustainable restoration and energy-efficient operation of CH buildings.

3. Knowledge Graph and Digital Twin Framework

In this Section we will analyse the nature of data (both IFC and weather data; IFC are the type of the files used in BIM representations) in Subsection 3.1. We will also give a high level description of the underlying KG (Subsection 3.2). The Section then proceed with the establishment of the problem that the DT is solving (Subsection 3.3). We conclude the section with a description of the use case upon which the pipeline will be tested. But first in Figure 1, we present the architecture of the pipeline. Based

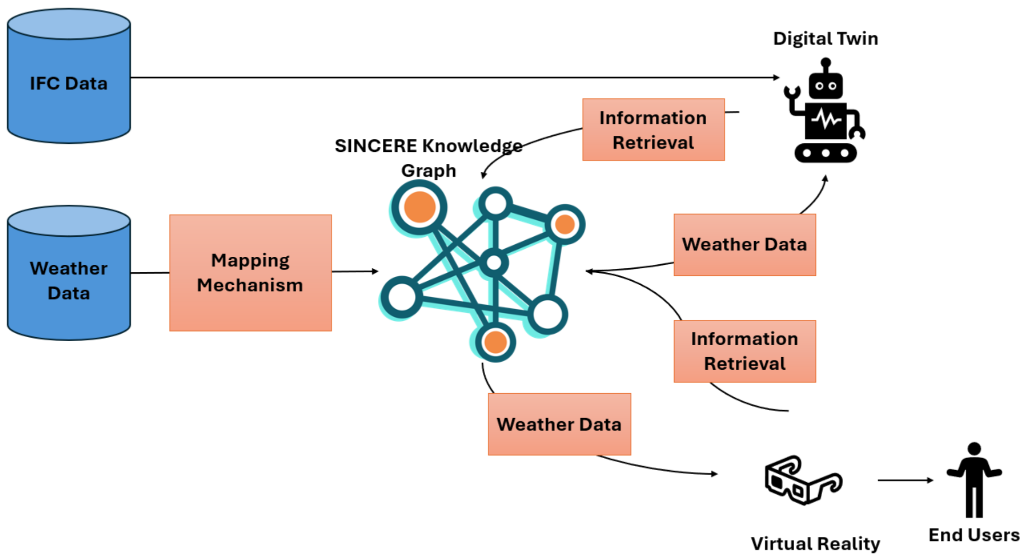

Figure 1: Architecture of the pipeline

on Figure 1, the weather data is translated into RDF using a mapping mechanism that converts the timeseries of weather data into RDF format. This data is then made accessible to the Digital Twin (DT) and end-users through a Virtual Reality tool. The Information Retrieval (IR) mechanism facilitates the distribution of data to both the DT and end-users via the POST/GET protocol. Through this protocol, a component or user provides input, which automatically generates a SPARQL query to retrieve the desired output. Additionally, Industry Foundation Classes (IFC) files are directly accessed by the DT, as they cannot be represented in the Knowledge Graph (KG). The source code for this tool is open-source[1].

3.1. Nature of Data

The real weather dataset for the Pilots consists of wind and temperature measurements across various building structures. We will take as running example the Spanish Pilot which has 4 buildings. The sensor distribution is as follows: the first building contains 8 sensors, the second has 16, while the third and fourth buildings have no sensors. Additionally, three sensors monitor cooling paint temperatures, and two capture weather data from solar panels.

Since the data collection methodology remains consistent across all cases, we focus on the first building, which includes 8 sensors. The dataset follows a time-series format, with measurements recorded at hourly intervals. Table 1 presents three consecutive time steps of recorded data.

The dataset structure consists of:

[1] https://github.com/valexande/sincere

Table 1
Real weather data for the first building of the Spanish Pilot. Values are rounded to two decimal places.

Location	2024-12-05 11:00:00	2024-12-05 12:00:00	2024-12-05 13:00:00
east_inside_lower	17.95	19.56	20.02
east_inside_upper	17.93	19.34	19.69
hf_east_inside_lower	9.26	8.18	5.37
hf_east_inside_upper	7.34	6.25	3.81
hf_west_inside_lower	10.22	9.52	7.56
hf_west_inside_upper	0.00	0.00	0.00
west_inside_lower	17.93	18.93	19.28
west_inside_upper rows	17.32	18.46	18.92

- **Time**: Timestamp in YYYY-MM-DD HH:MM:SS format.
- **Temperature Measurements**: Collected at different building locations (e.g., East Inside Lower, West Inside Upper).
- **Wind Speed Measurements**: Recorded at corresponding locations (e.g., HF East Inside Lower, HF West Inside Upper).

This dataset serves as a foundation for analyzing the real impact of weather on building performance, aiding in the development of energy optimization strategies within the Spanish Pilot. It is important to note that both past and future weather data were predicted, as real-life sensor measurements were unavailable. The predictions were generated using weather models trained on data such as temperature, wind speed, humidity, and solar radiation from locations with similar climates. For example, the Spanish Pilot is located in Algete, a village outside Madrid. Since no direct weather data was available for Algete, the prediction model was trained using data from Madrid. Lastly, while the dataset includes additional measurements such as solar radiation, total humidity, wind speed, snow, and rain percentage, these were omitted due to space constraints.

3.2. SINCERE Knowledge Graph

For each step of the timeseries a URI in the class Observation is created, then for each different measurement a datatype property is attached to the Observation class. For instance, in the example from SubSection 3.1 we will have 8 datatype properties to represent the data in east_inside_lower, east_inside_upper, hf_east_inside_lower, hf_east_inside_upper, hf_west_inside_lower, hf_west_inside_upper, west_inside_lower, west_inside_upper rows. The naming of the datatype properties follows the naming of the rows. Equivalently, for the second building where we have 16 sensors we will have 16 different datatypes. The measuring units is also included in the values of each property. The KG is composed from other class and properties for each measurement, but we leave here only a high level analysis of how the data is represented in the KG.

As mentioned in the introduction, the KG also leverages semantic relationships within the data to provide a more accurate estimation of the actual temperature at a specific location within the building. Before passing the data to the DT, the KG consolidates additional measurements—such as humidity at the sensor's location, wind speed, and solar radiation—to refine the temperature readings. The influence of these factors is calculated using physical equations from relevant literature: solar radiation is accounted for using the model in [10], humidity is incorporated based on [11], and wind speed effects are derived from [12].

3.3. Digital Twin

Upon receiving the IFC model of the location and weather data, the DT must solve an optimization problem. Its primary objective is to maintain room temperatures at a specified level, as defined by domain experts, to protect the materials of the CH site. Additionally, the DT seeks to minimize energy consumption from Air Conditioning (AC) systems, thereby reducing the building's overall carbon footprint.

To achieve this, the DT evaluates various materials that can be applied to the internal and external surfaces of the building to optimize hourly AC energy consumption. By modifying the virtual counterpart of the building, the DT enables testing in a risk-free environment, reducing both costs and potential structural risks associated with real-world material trials. The materials under consideration must not compromise the integrity of the existing building structure while being designed to absorb heat and enhance passive warming.

Ultimately, the DT recommends the best-performing material—one that minimizes AC energy consumption while maintaining a stable indoor temperature—ensuring both energy efficiency and the preservation of the heritage site.

3.4. Reducing the Carbon Footprint and Increasing Energy Performance

Our pipeline will be tested in four different locations: a controlled environment consisting of four small buildings in Algete-Spain (see Figure 2), a CH site in Rhodes-Greece, an industrial site in Ostrava-Czechia, and another CH site in Holon-Israel.

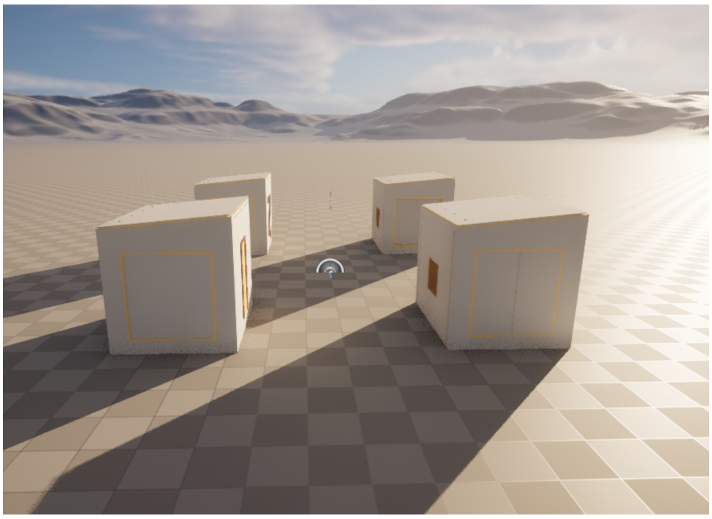

Figure 2: Algete Spain pilot in BIM representation

All of the aforementioned sites will be equipped with the necessary sensors to store their data in the KG. Their virtual counterparts will be provided to the DT component through BIM representations. Additionally, various materials will be tested based on the climate characteristics of each region. For instance, the warm, dry climate of Israel contrasts with the warm, humid conditions of Rhodes and the colder climate in Ostrava, each requiring different materials to achieve the goal outlined in Subsection 3.3.

4. Conclusion

The integration of KGs, DTs, and BIM in the proposed pipeline demonstrates a promising approach to improving the sustainability and energy efficiency of historic buildings. By leveraging real-time and predictive weather data, the system enables dynamic decision-making for selecting restoration materials that optimize energy consumption while preserving the architectural integrity of CH structures. The use of advanced weather prediction models and simulation techniques offers a data-driven methodology for addressing climate change challenges in built heritage conservation. However, certain limitations must be acknowledged. The accuracy of the predictive models remains a concern, as their reliability depends on the quality and availability of historical data. Additionally, the methodology has been tested exclusively on CH buildings, and adaptations may be necessary for broader applications, such as industrial or commercial buildings.

The results obtained from the tested pilot sites in Spain, Greece, Czechia, and Israel validate the system's capability in diverse climatic conditions. However, further testing and refinement are required to enhance its robustness. A key area for future work involves improving the weather prediction models by incorporating more extensive datasets and advanced machine learning techniques. Moreover, expanding the KG framework to integrate additional environmental factors, such as air pollution and moisture infiltration, could further refine the decision-making process for material selection. Another promising direction is the automation of material recommendations by incorporating AI-driven optimization techniques within the DT framework, ensuring minimal human intervention while maximizing sustainability outcomes.

Looking forward, the project aims to enhance the interoperability of the proposed system by developing standardized APIs that facilitate integration with existing smart city infrastructures. Additionally, stakeholder engagement and policy integration will be essential in ensuring widespread adoption, particularly among heritage conservation agencies and urban planners. The long-term goal is to establish a comprehensive, AI-enhanced decision-support system that not only aids in historic building preservation but also contributes significantly to global net-zero-carbon initiatives. Through continuous refinement and real-world validation, this research paves the way for a scalable and impactful solution to the pressing challenge of sustainable heritage conservation in the face of climate change.

5. Acknowledgments

This work has been funded by SINCERE Horizon Europe project, grant agreement number 101123293.

Declaration on Generative AI

During the preparation of this work, the authors used Grammarly in order to grammar and spell check, and improve the text readability. After using the tool, the authors reviewed and edited the content as needed to take full responsibility for the publication's content.

References

[1] Murphy, M., McGovern, E., & Pavia, S. (2009). Historic building information modelling (HBIM). Structural Survey, 27(4), 311-327.

[2] Banfi, F. (2017). BIM orientation: grades of generation and information for different type of analysis and management process. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 42, 57.

[3] Grieves, M. (2017). Digital Twin: Manufacturing Excellence through Virtual Factory Replication. Manufacturing Leadership Journal.

[4] Boje, C., Guerriero, A., Kubicki, S., & Rezgui, Y. (2020). Towards a semantic Construction Digital Twin: Directions for future research. Automation in Construction, 114, 103179.

[5] Li, C., Xie, H., Lin, J., Xue, F., & Yang, L. (2020). A review of the application of digital twin in construction. IEEE Access, 8, 184318-184333.

[6] Lu, Qiuchen and Parlikad, Ajith Kumar and Woodall, Philip and Don Ranasinghe, Gishan and Xie, Xiang and Liang, Zhenglin and Konstantinou, Eirini and Heaton, James and Schooling, Jennifer (2020). Developing a digital twin at building and city levels: Case study of West Cambridge campus. Journal of Management in Engineering, 36(3), 05020004.

[7] Hogan, Aidan and Blomqvist, Eva and Cochez, Michael and d'Amato, Claudia and Melo, Gerard De and Gutierrez, Claudio and Kirrane, Sabrina and Gayo, José Emilio Labra and Navigli, Roberto and Neumaier, Sebastian and others (2021). Knowledge graphs. ACM Computing Surveys (CSUR), 54(4), 1-37.

[8] Chen, Y., Liu, Y., Mao, C., & Wang, H. (2020). Knowledge graph for historical building information modeling. Automation in Construction, 114, 103187.
[9] Garijo, D., Poveda-Villalón, M., & Corcho, O. (2022). Knowledge graphs for sustainable development: Challenges and opportunities. Journal of Web Semantics, 72, 100679.
[10] La Gennusa, M., Nucara, A., Pietrafesa, M., & Rizzo, G. (2007). A model for managing and evaluating solar radiation for indoor thermal comfort. Solar Energy, 81(5), 594-606.
[11] Arundel, A. V., Sterling, E. M., Biggin, J. H., & Sterling, T. D. (1986). Indirect health effects of relative humidity in indoor environments. Environmental health perspectives, 65, 351-361.
[12] Oh, W., & Kato, S. (2018). The effect of airspeed and wind direction on human's thermal conditions and air distribution around the body. Building and Environment, 141, 103-116.