

A GRAPH BASED FRAMEWORK TO SUPPORT DATA-DRIVEN URBAN BUILDING ENERGY SIMULATIONS

Usman Ali¹, Sobia Bano¹, Cathal Hoare¹, Muhammad Haris Shamsi²,
Divyanshu Sood¹ and James O'Donnell¹

¹School of Mechanical and Materials Engineering and UCD Energy Institute, UCD, Dublin, Ireland

²VITO, Hasselt, Belgium

Abstract

Urban planners and energy policymakers increasingly focus on sustainable urban development and the challenges of analyzing complex urban energy systems. Current models often lack the integration of diverse urban datasets and do not adequately address the dynamic nature of urban energy demands. This study proposes a data-driven framework that involves data collection and preprocessing, building archetypes, machine learning modeling, and parametric simulation. The novel contribution of this research lies in defining the scope, processes, information, data, and relationships for the ontology of urban building energy modeling, employing a graph-based approach for complex data integration. The proposed methodology is tested in residential buildings in Dublin City to examine and compare the modeling results. The study concludes that the proposed model offers a more comprehensive and adaptable approach to urban energy analysis compared to traditional methods. Furthermore, the study helps stakeholders by providing a scalable and flexible modeling framework for urban energy analysis.

Introduction

In the era of rapid urbanization, sustainable development in cities has become a crucial global challenge. The European Union has established a robust legislative framework to enhance sustainable planning and improve the energy performance of buildings. This framework is underpinned by two key directives, the Energy Performance of Buildings Directive EU/2010/31 and the Energy Efficiency Directive EU/2023/1791, both revised in 2023. These directives lead EU member states to adopt policies that aim to achieve a highly energy-efficient and decarbonized building stock by 2050 (EU-Energy, 2023).

Efficient and sustainable urban environments are critical for identifying scalable energy conservation strategies. A promising approach involves analyzing building energy performance data at an urban scale using a data-driven approach. However, the available urban data are often sparse, inconsistent, and lacking in diversity and heterogeneity. Despite these challenges, the last few decades have seen significant advancements in data-driven modeling, particularly in using sparse data to predict and estimate building energy usage. Nevertheless, a gap remains in these studies, primarily due to their focus on prediction and estimation without adequately integrating complex, multidimensional data sets. Therefore, a more comprehensive approach is essential for integrating complex urban building

energy dynamics data to improve urban-scale modeling.

Urban planners and energy policymakers increasingly focus on the complexities of urban energy systems, driven by the crucial need to balance environmental sustainability with the growing energy demands of expanding urban populations. However, there is a lack of comprehensive models that effectively integrate diverse urban data sets to analyze urban energy systems. Current approaches often fail to address the dynamic and complex nature of urban energy demands. There is a growing recognition of the gap in the literature, where traditional models are insufficient to capture multifaceted interactions within urban environments. This gap necessitates a novel approach that can holistically encapsulate the intricate energy dynamics of cities.

This study introduces an innovative data-driven framework for urban building energy modeling, taking advantage of ontology and graph-based approaches. The ontology provides a structured representation of domain knowledge, allowing a clearer definition of concepts, attributes, and relationships (Poveda-Villalón et al., 2022; Curry et al., 2013; Hoare et al., 2022). Combining the ontology with a graph database enables creating of a comprehensive knowledge representation system that can efficiently handle diverse and interconnected data, promoting a more holistic understanding of urban energy systems.

Ontology and graph-based techniques are widely used in the field of Building Information Modeling (BIM), linked building data, urban planning, and energy management (Pritoni et al., 2021; Terkaj et al., 2017; Costin and Pauwels, 2022; Li et al., 2019). Kapsalis et al. (2022) use a graph analysis approach to efficiently query and analyze energy efficiency certificates. Wu et al. (2021) propose an ontology modeling solution for managing decentralized data for household energy systems. Van Dam and Keirstead (2010) initialize a model of an urban energy system built on an ontology using energy transformation. Wu et al. (2022) propose an ontology-based framework that can integrate data to build energy simulations from different data sources in the Operation Phase. Baumgärtel et al. (2014) used an ontology framework to assess the application of building performance regulations in design and operation. Daneshfar et al. (2022) proposed an ontology to represent geospatial data to support building renovation to collect data in IFC and CityGML format. Zadeh et al. (2019) developed a hybrid information infrastructure by integrating building design data in ifcXML format and 3D neighborhood models in CityGML format. Shi et al.

(2023) proposed a methodology based on ontology to create a digital twin city information model by integrating BIM, GIS, and IoT technologies. However, existing studies focus mainly on urban building energy management and building information modeling, but more research is needed in the field of urban building energy modeling. Furthermore, available urban building energy models typically do not emphasize data integration, flexibility, scalability, and performance, which are critical to adapting and managing the complexities of urban energy systems.

The novelty of this research lies in defining the scope, processes, information, data, and relationships for an urban building energy modeling ontology. The application of such an ontology uses a graph-based approach to assimilate and analyze complex urban energy data, enabling a more holistic understanding of urban energy dynamics. Furthermore, the framework enriches the predicted machine learning-based data of complex energy modelling scenarios, an aspect rarely explored in existing studies. This integration allows integration of diverse datasets, ranging from building stock to energy consumption patterns, and fosters a deeper understanding of the interdependent factors that influence urban energy systems. This study aims to offer a more comprehensive and adaptable approach to urban energy analysis, surpassing the limitations of traditional methods.

The article is organized as follows: Section 2 provides a detailed discussion of the methodology devised for urban energy modeling of residential buildings. Section 3 discusses the case study of Irish building stock. Finally, conclusions are discussed in Section 4.

Methodology

Modeling the energy performance of buildings on a large urban scale presents a formidable challenge for urban planners and policymakers. Accurate prediction of energy consumption and identification of energy efficiency opportunities are essential to promote the sustainable development of cities.

Therefore, this study proposes a framework for modeling the energy performance of buildings on an urban scale, which begins with the development of an ontology and its implementation in a graph database. This lays the founda-

tional framework for representing and analyzing the complex interrelations of urban energy performance (Figure 1). The process involves comprehensive data collection and preprocessing to gather and refine building-related data. Subsequently, building archetypes are developed to define representative models of buildings, capturing the diversity of the urban built environment. These archetypes serve as the basis for parametric simulations, exploring the energy performance under various scenarios and conditions and generating a synthetic dataset for analysis. Machine learning models are then employed to predict energy performance across the urban scale, leveraging the insights gained from synthetic and real-world data to identify opportunities for energy efficiency improvements. Finally, this structured approach facilitates querying and data analysis, providing a detailed understanding of urban building energy performance. This holistic approach integrates sophisticated data analysis and modeling techniques to offer actionable insights for promoting sustainable urban development.

Urban Building Energy Modeling Processes

The proposed ontology for urban building energy modeling includes building archetypes and parametric simulation data, which are used to generate synthetic data (Figure 3). Ontology also incorporates machine learning modeling results, focusing on predicting the energy performance of buildings based on their respective archetypes. The proposed ontology includes entities such as *BuildingArchetypes*, *BuildingParameters*, *ConstructionTemplates*, *EnergyUsage*, and *MLPredictedEnergyPerformance*. As a result, this interconnected data enrichment creates a robust platform for urban modeling. This improves understanding of the energy dynamics of urban buildings and enables the development of targeted strategies to improve energy efficiency. The proposed ontology mapping on desired scenarios requires data collection and preprocessing, archetype development, parametric simulation, and machine learning modeling processes. These processes are designed to map complex urban data and scenarios from existing buildings onto ontologies and predict building energy performance.

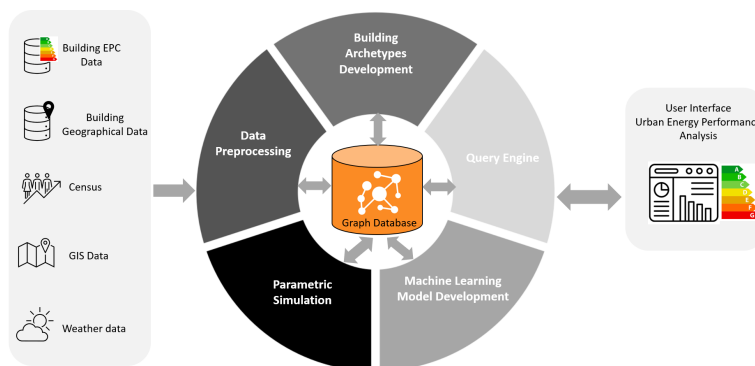


Figure 1: Methodology for data-driven graph-based urban building energy modeling

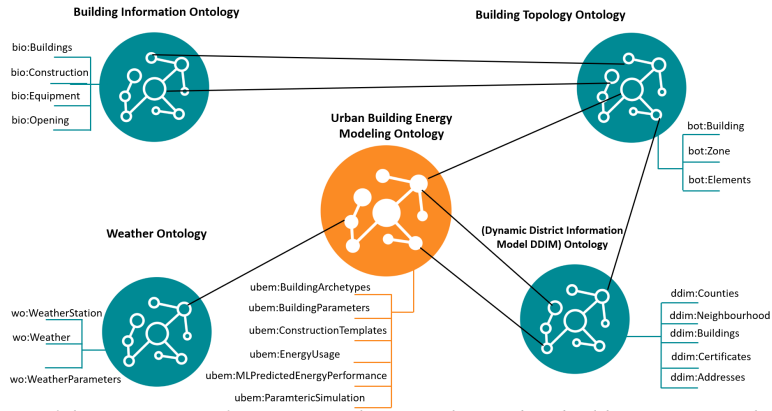


Figure 2: A conceptual view of the integration of existing ontologies with an urban building energy modeling ontology with required entities

Graph Database Development

A graph database is developed to efficiently manage and store complex interconnected data. This database enables the representation of relationships between various data points, such as the link between building types and energy usage patterns, facilitating more detailed analysis and insights (Curry et al., 2013; Zhu et al., 2022). The commonly used graph databases are Neo4j, Amazon Neptune, and OrientDB. A graphical representation of the complex graph model can be presented using the ontology (Hoare et al., 2022).

Ontology refers to a structured representation of knowledge that defines concepts, relationships, and properties within a specific domain, such as urban energy systems (Costin and Pauwels, 2022). These ontologies are represented using the Web Ontology Language (OWL). The proposed ontology for urban building energy modeling has been developed by integrating existing ontologies into the building domain, such as the Building Information Ontology, Building Topology Ontology, Weather Ontology, and Dynamic District Information Model Ontology (Figure 2). Building Information Ontology (BIO) provides a range of defined classes, axioms, and data types for reuse, including Building, BuildingElement, and BuildingParameter (TUWien, 2024a). Building Topology Ontology (BOT) is a minimal ontology to describe the core topological concepts of a building (Rasmussen et al., 2021). To represent weather-related information for a location, the specific Weather Ontology (WO) provides reusable patterns and terms (TUWien, 2024b). The Dynamic District Information Model (DDIM) offers a national-scale digital twin for domestic building stock, complete with geographic information (Hoare et al., 2022).

Data Collection

The ontology creation step starts with collecting available raw homogeneous data sources, including building energy performance certificate (EPC) data, geographic data, census information, GIS, and weather data. These data are gathered from various sources, including national databases and urban planning departments. The data are

carefully selected to ensure relevance and accuracy for urban energy modeling.

Data Preprocessing

Once collected, the data undergo a rigorous pre-processing phase. This process involves cleaning the data and handling missing values to ensure consistency. Pre-processing also includes categorizing buildings based on characteristics such as age, region, and type of usage, which is crucial for developing accurate building archetypes. The pre-processed data are mapped to existing BIO, BOT, DDIM, and WO, further supporting the urban building energy ontology.

Building Archetypes Development

Building archetypes are developed to represent different categories of buildings with similar characteristics within the urban scale. Each building archetype serves as a core entity (*ubem:BuildingArchetypes*) for graph-based integration of urban buildings and requires specified features, such as geometry, layout, construction materials, insulation levels, and typical energy use patterns. These data can be extracted from existing ontologies, namely the BOT and the BIO. Similarly, the DDIM ontology provides all the geographical information for the building archetype. These archetypes serve as foundational elements for a parametric simulation framework that models energy performance in various types of buildings. In addition, these archetypes are essential for urban modeling because they encapsulate common building characteristics that can be used to generalize energy performance assessments.

Parameters Simulation

Parametric simulation generates synthetic data to simulate various energy consumption and performance scenarios to find the optimal solution, especially when a sparse data set is available for energy modeling (Ali et al., 2024). Parametric simulation uses developed archetypes and also sources data from the weather ontology for simulation. It stores all input and output results in a graph-based structure using building parameters (*ubem:BuildingParameters*), Construction (*ubem:ConstructionTem-*

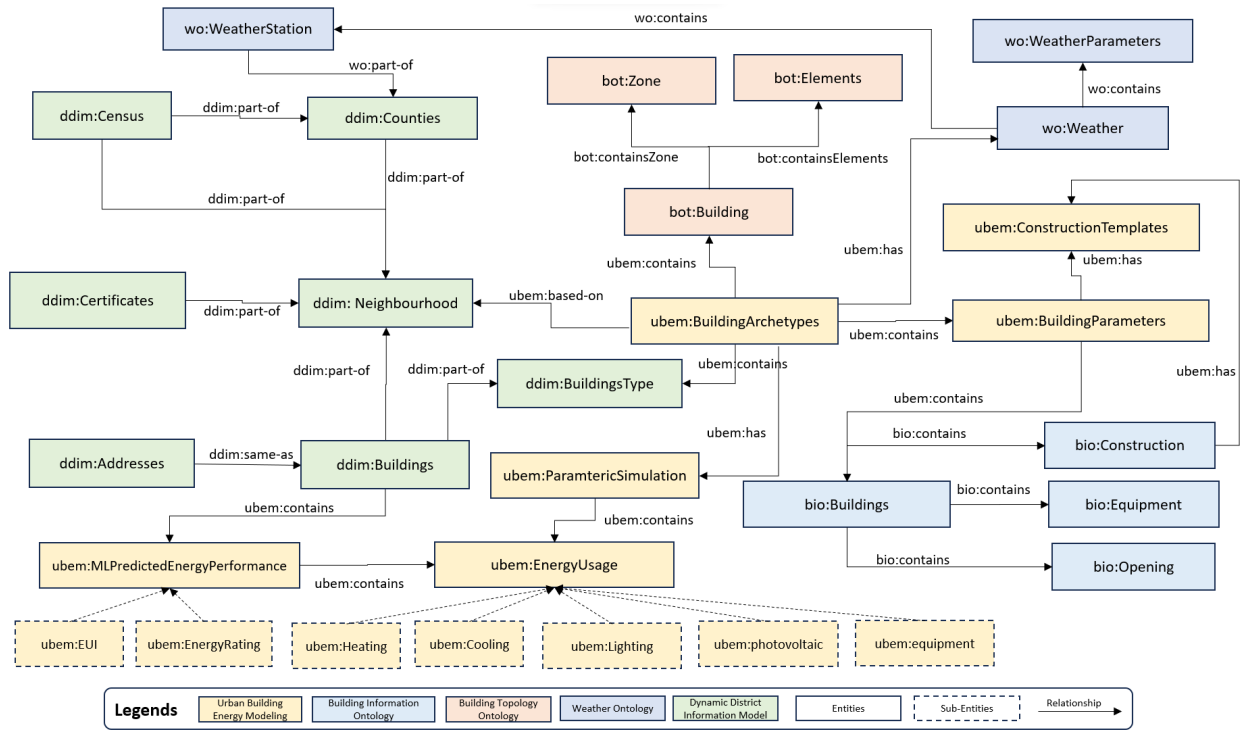


Figure 3: Graph structure for urban building energy modeling for residential building stock

plates), and the energy usage (*ubem:EnergyUsage*) entity. (*ubem:ParametricSimulation*) entity estimates are determined using thermal modeling to assess heating, lighting, and water demands, and employing an energy simulation engine to evaluate the overall energy performance of the building under different conditions. A parametric tool performs numerous simulations using a simulation model (Ali et al., 2024) to perform complex parametric simulations involving multiple parameters. In this paper, JEPlus serves as a parametric tool for energy simulations. Additionally, JEPlus uses EnergyPlus for simulations, integrating various parameter values together with weather data and construction templates. However, due to the complex nature of the numerous parameters involved, generating simulated data for all parameters becomes nearly impossible. Therefore, synthetic data are generated using sampling methods such as Simple Random Sampling (SRS) and Latin Hypercube Sampling (LHS). These methods help generate the desired sample data that includes combinations of all parameters. Parametrically simulated data produce synthetic data sets stored back in the graph database and then used as input for developing machine learning models.

Machine Learning Modeling

Data-driven Machine-Learning (ML) models have been designed to enrich the predicted results for urban energy use scenarios in graph databases. The workflow involves formulating regression ML models to store predicted results in the energy performance entity of the building, specifically in terms of Energy Use Intensity (EUI) and energy rating. The model is trained using the data stored in the graph database. Generally, the models developed

include the process of splitting data into training and testing sets, followed by the application of regression algorithms based on performance indices (Ali et al., 2024). These models are continuously refined to improve accuracy and adaptability to evolving urban dynamics. Furthermore, the models predict the intricate characteristics of buildings on an urban scale and store these predicted data in a graph database using *ubem:MLPredictedEnergyPerformance* entity for further analysis. The enrichment of ML-based building energy performance in graph-based databases offers dynamic, context-aware insights, significantly enhancing the precision and relevance of urban building energy modeling.

Query Engine

Developing a query engine enables efficient retrieval and data analysis from the graph database developed based on the proposed ontology. This engine supports complex queries, allowing users to extract specific insights, such as determining the influence of particular building characteristics on energy consumption or identifying potential areas for energy efficiency improvements. The most common graph database query languages are Cypher, GraphQL, and SPARQL (Kapsalis et al., 2022; Zhu et al., 2022; Hoare et al., 2022). Cypher is one of the most widely used query languages for graph databases, especially in the context of Neo4j. GraphQL is a query language primarily designed for APIs, but it can also be employed to query data from graph databases. On the other hand, SPARQL is a query language specifically used for querying RDF (Resource Description Framework) data. The proposed ontology, coupled with graph-based integration for urban

Table 1: Data requirements and associated data sources for Dublin graph-based database

Ontology	Prefix	Irish Data Source	Publisher
Building Information Ontology	bio	Irish EPC (BER) Database	SEAI
Weather Ontology	wo	Dublin EPW File	EnergyPlus
Building Topology Ontology	bot	Construction Templates, GeoIntel	An Post/ Ordnance Survey, DesignBuilder
Dynamic District Information Model	ddim	GeoDirectory, Irish Census database	An Post/ Ordnance Survey Ireland, Central Statistics Office
Urban Building Energy Modeling	ubem	Irish GIS database	Central Statistics Office

building energy modeling, can be further examined using a query engine based on data integration, flexibility, scalability, and performance. This study examines the query engine with and without machine learning-based data enrichment for estimating urban building energy performance.

energy. Adopting an ontology-driven methodology supported by Neo4j graph database technology transforms the management and analysis of Dublin’s residential building stock. Furthermore, this study helps in detail urban energy insights through a more structured, interconnected, and efficient data framework than conventional approaches.

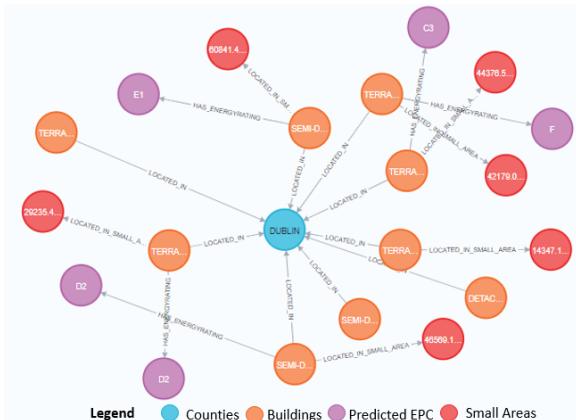


Figure 4: Visual graph representation of relationship complex urban building energy modeling for Dublin city

Irish Graph Database Development

This study uses a Neo4j graph database to implement the proposed urban building energy modeling ontology. Neo4j efficiently stores nodes and relationships, significantly streamlining complex data management. Using Neo4j’s capabilities, the study effectively represents and interlinks various urban data. The interconnected data enables comprehensive analysis of urban energy, facilitating the identification of patterns and insights that are difficult to identify with traditional methods due to complex relationships and interdependencies at an urban scale. The graph-based structure of Neo4j also allows high-performance querying and data retrieval, making it an ideal platform for handling complex queries essential for comprehensive urban energy modeling and analysis.

Urban Energy Performance Analysis

Finally, the methodology culminates in a comprehensive analysis of urban energy performance. This involves using the models and tools developed to evaluate and interpret buildings’ energy efficiency and the overall energy dynamics of the urban area. Graphical or GIS-based modeling results can be integrated into the analysis using data stored in the DDIM ontology for geographical information. The analysis provides valuable information on current performance levels and identifies opportunities for optimization and improvement.

Gathering data on an urban scale for a building stock is a challenging task, as individual building information is often limited and sparse. The data collection process for this study involved acquiring raw building data from multiple sources to map them to the proposed or existing ontology. The case study includes data such as building energy performance certificates, building geographical data, census information, GIS, and weather data (Table 1).

The study aims to provide a robust and comprehensive framework for urban building energy modeling, providing a deeper understanding of urban energy dynamics and paving the way for more sustainable urban development.

In Ireland, the data set for EPC (also known as the Building Energy Rating (BER) certificate) of Irish residential stock represents a comprehensive measurement of the building stock, which includes more than 200 characteristics of the buildings. These characteristics include the building fabric, heating systems, estimated end use, CO₂ emissions, and estimated delivered and primary energy consumption. The Irish EPC dataset contained approximately 1.1 million residential buildings, with a substantial number of building ratings that fall within the C1 to D2 range (SEAI, 2023). EPC data assists in retrieving building-related data for archetype development, based on building topology and building information ontology requirements.

Case Study

The primary objective of this case study is to test the proposed methodology on the residential building stock of Dublin, Ireland. This study demonstrates the effectiveness of an ontology-driven approach in the management and analysis of complex urban building data. The application of this data-driven framework to buildings in Dublin has revealed significant insights into the dynamics of urban

The Irish census provides spatial data on various scales and the number of buildings in each geographic area, such as small areas (neighborhoods), and counties (CSO, 2022). Furthermore, the GeoDirectory database, updated by An

Post (Irish Postal Service) and Ordnance Survey Ireland, offers statistical and geographical information on the entire Irish building stock. The Q2 2023 GeoDirectory report includes geocoded addresses of 2.1 million residential buildings (GeoDirectory, 2023; OSI, 2023). This study focuses on Dublin City, and the EPC data is available only for about 54% of the residential building stock in Dublin City. GeoDirectory is a useful resource for urban energy modeling and is required for the DDIM Ontology to analyze geographical aspects of modeling results. Furthermore, Dublin weather data is used to evaluate the impact of weather conditions on energy performance based on weather ontology requirements. This study integrates sparse and heterogeneous urban building data using a graph-based approach. Then, machine learning algorithms are applied to this integrated data set to predict the energy rating of the remaining 46% building stock.

After the initial data collection process, all data undergo a rigorous preprocessing phase. This phase involves cleaning the data and addressing missing values to ensure consistency. The preprocessed data mapped to the ontology are then used to create a Neo4j graph-based database. This step is crucial as it involves removing all irrelevant data and including only relevant data, which is essential to improve the performance of the query engine and the execution of complex queries for urban-scale modeling. For example, the Irish EPC database contains more than 200 features, and this process selects important features for building energy modeling based on existing studies (Ali et al., 2024).

This study uses parametrically simulated data comprising 1 million entries, generated using four residential archetypes such as semi-detached, detached, terraced, and bungalow. These archetypes were developed using available Irish building data based on data requirements of BIO, BOT, and DIMM ontologies. These parametric simulated data are created using DesignBuilder construction templates for building parameter mapping. Parametric data are stored in Neo4j and are further used for data-driven machine-learning modeling to predict the energy performance of urban buildings. Neo4j's graph database structure allows for the seamless integration of diverse Dublin data sources, providing a more comprehensive view of urban energy systems compared to traditional data approaches. Neo4j can help improve the generation of models for energy performance for urban buildings due to the richness of data and relationships available in the graph. The data are then partitioned into two subsets to create training and testing datasets, employing a cross-validation algorithm. These data sets were trained and evaluated using regression models (eXtreme Gradient Boosting) proposed in the existing study (Ali et al., 2024). Finally, the machine learning-based enrichment results are compared with those of traditional simple queries for building analysis.

```
MATCH (b:Buildings)-[:part-of]->(sa:SmallAreas),
      (b)-[:part-of]->(c:Counties {COUNTY: 'DUBLIN'})
WHERE b.BER_ENERGY_MEDIAN > 300
RETURN sa, collect(b) as Buildings
```

Figure 5: A complex query that finds buildings in small areas within Dublin County with an energy performance greater than 300 kWh/m²/yr EUI.

Dublin Energy Performance Analysis

The results show that the Neo4j graph-based database efficiently manages and stores complex data using the proposed ontology for Dublin, demonstrating its interoperability in combining various formats, including spatial and building energy data. The database stores data, including parametric simulated pre-processed data and predicted energy performance data for buildings based on their respective archetypes in Dublin (Figure 4). These interconnected data sets establish a robust foundation for urban modeling, deepening our understanding of the energy dynamics within urban buildings. Furthermore, the graph-based database developed for Dublin buildings exhibits sufficient scalability and flexibility, making it suitable for handling large and complex datasets. This accommodates the ever-evolving nature of urban energy systems and supports various urban planning scenarios.

The Cypher query offers a valuable tool for urban studies, specifically in the context of energy consumption within building infrastructures. For instance, the Cypher query helps to analyze energy efficiency within Dublin County's building infrastructure by identifying buildings with high energy use intensity, specifically those with a median energy performance above 300 kWh/m²/yr EUI (Figure 5). Similarly, the Cypher query can help with urban energy modeling in Neo4j, which is a vital tool input for the simulation of building energy performance within an urban context (Figure 6). The query gets the building parameters from BuildingParameter nodes and inputs them to Simulation nodes. With these parameters, the query sets the stage for parametric simulations that calculate energy performance metrics such as Energy Use Intensity (EUI), heating, lighting, and water usage. This process enables the detailed examination of building performances and facilitates a broader analysis of urban energy consumption patterns. The ability to dynamically adjust and simulate various building parameters provides a data-driven approach for stakeholders.

The developed graph can also help with the breakdown of buildings according to their Energy Rating from "A1" to "G." (Figure 7) The results showed a diverse range of energy efficiencies among the buildings, with the count of buildings for each energy rating varying significantly. The highest number of buildings fell into the "C2" rating, with 24,539 buildings indicating moderate energy efficiency. In contrast, the "A1" rating has the fewest buildings, with only 1,380, signifying the highest energy efficiency. Other notable findings include many buildings in the "D1" and "C3" ratings, with 24,307 and 23,252 buildings, respec-

tively. This distribution highlights the poor energy efficiency in buildings across Dublin, providing insights into the current state of energy performance and potential areas for improvement in building energy standards.

```

MATCH (bp:BuildingParameter), (s:Simulation {id: bp.id})
SET s += {
  floorUValue: bp.`floor-uvalue`,
  doorUValue: bp.`door-uvalue`,
  roofUValue: bp.`roof-uvalue`,
  windowUValue: bp.`window-uvalue`,
  wallUValue: bp.`wall-uvalue`,
  ...
}
WITH s
RETURN s.id AS SimulationID,
  s.EUI AS EUI,
  s.EnergyRating AS EnergyRating,
  s.HeatingUsage AS HeatingUsage,
  s.LightingUsage AS LightingUsage,
  s.WaterUsage AS WaterUsage

```

Figure 6: A complex query that get building parameters for parametric simulation for energy performance calculation

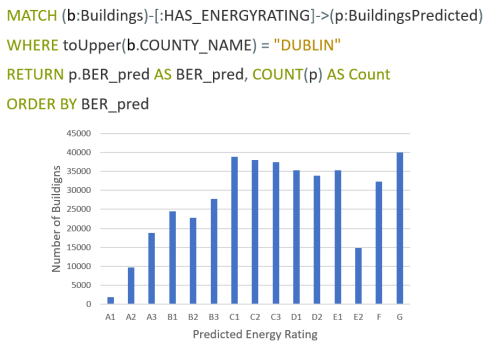


Figure 7: A cypher query show the distribution of predicted building energy ratings in Dublin

Finally, the developed graph database allows stakeholders to formulate targeted strategies to improve energy efficiency in Dublin, including existing and predicted data, for use in future analysis. Cypher graph queries are used to analyze data within the database, enabling complex analysis. The results can be visualized on a GIS map, highlighting buildings with poor energy performance in a specific Dublin county area, based on Cypher graph queries (Figure 8). This visualization helps decision-makers identify key areas for energy efficiency improvements and informs policy decisions to promote sustainability and reduce energy consumption in the region. Moreover, the flexibility of the database in handling various data formats, including spatial and building energy data, enhances its utility for comprehensive urban planning and energy management initiatives.

Conclusions

This study underscores the growing importance of sustainable urban development and the need for an advanced framework to analyze complex urban energy data. The motivation behind this research was to address the

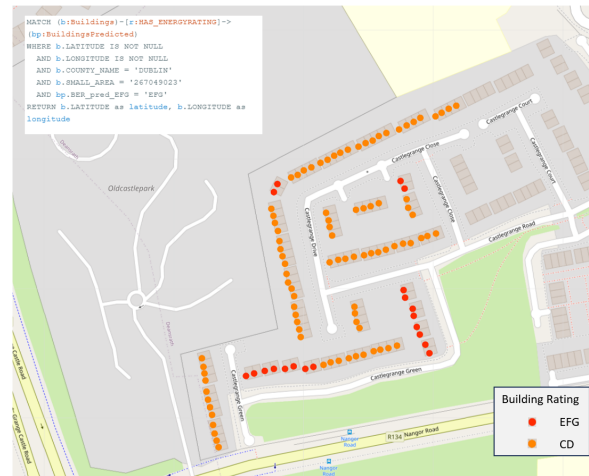


Figure 8: Visual map of buildings with poor energy performance in a specific Dublin county area based on cypher graph query

limitations of current models, which often need help integrating diverse urban data sets and addressing the dynamic nature of urban energy demands.

The key contributions of this research lie in developing a novel data-driven framework. This framework, which encompasses data collection, preprocessing, building archetypes, parametric simulation, and machine learning modeling, leads to the creation of a graph database with ontology and query engine. This approach enables a more holistic understanding of urban energy dynamics while accommodating diverse datasets. An ontology for urban building energy modeling facilitates the integration of complex urban energy data, ultimately offering a more comprehensive and adaptable approach to urban energy analysis compared to traditional methods. This study empowers various stakeholders to analyze and predict complex energy scenarios, thus supporting the creation of more sustainable and energy-efficient urban environments. The application of the proposed framework could be expanded to include commercial buildings and energy suppliers to assess its applicability and effectiveness in diverse contexts. In general, this research lays the foundation for future advancements in sustainable urban development and energy policy, offering a promising direction for further exploration and innovation in this critical area.

Acknowledgments

This publication has emanated from the US-Ireland R&D Partnership supported by the Science Foundation Ireland through 20/US/3695, the U.S. National Science Foundation through Award Number 2217410, and the Department for the Economy in Northern Ireland through USI 167. We would also acknowledge NexSys project supported by the Science Foundation Ireland through Award Number SFI/21/SPP/3756. The opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funding agencies.

References

- Ali, U., Bano, S., Shamsi, M. H., Sood, D., Hoare, C., Zuo, W., Hewitt, N., and O'Donnell, J. (2024). Urban building energy performance prediction and retrofit analysis using data-driven machine learning approach. *Energy and Buildings*, 303:113768.
- Baumgärtel, K., Kadolsky, M., and Scherer, R. J. (2014). An ontology framework for improving building energy performance by utilizing energy saving regulations. In *Proceedings of European conference on product and process modelling*, pages 519–526.
- Costin, A. and Pauwels, P. (2022). Building information modeling and ontologies: Overview of shared representations. *Research Companion to Building Information Modeling*, 12:12–34.
- CSO (2022). *Census of Population 2022 - Profile 1 Housing in Ireland* by Central Statistics Office. <https://www.cso.ie>.
- Curry, E., O'Donnell, J., Corry, E., Hasan, S., Keane, M., and O'Riain, S. (2013). Linking building data in the cloud: Integrating cross-domain building data using linked data. *Advanced Engineering Informatics*, 27(2):206–219.
- Daneshfar, M., Hartmann, T., and Rabe, J. (2022). An ontology to represent geospatial data to support building renovation. *Advanced Engineering Informatics*, 52:101591.
- EU-Energy (2023). *Energy for Europe* by European Commission. https://energy.ec.europa.eu/index_en. [Online; accessed 08-Feb-2024].
- GeoDirectory (2023). *Geodirectory technical guide, an post and ordnance survey Ireland*.
- Hoare, C., Aghamolaei, R., Lynch, M., Gaur, A., and O'Donnell, J. (2022). A linked data approach to multi-scale energy modelling. *Advanced Engineering Informatics*, 54:101719.
- Kapsalis, P., Kormpakis, G., Alexakis, K., and Askounis, D. (2022). Leveraging graph analytics for energy efficiency certificates. *Energies*, 15(4):1500.
- Li, Y., García-Castro, R., Mihindukulasooriya, N., O'Donnell, J., and Vega-Sánchez, S. (2019). Enhancing energy management at district and building levels via an em-kpi ontology. *Automation in Construction*, 99:152–167.
- OSI (2023). *Ordnance Survey Ireland*. <https://www.osi.ie>.
- Poveda-Villalón, M., Fernández-Izquierdo, A., Fernández-López, M., and García-Castro, R. (2022). Lot: An industrial oriented ontology engineering framework. *Engineering Applications of Artificial Intelligence*, 111:104755.
- Pritoni, M., Paine, D., Fierro, G., Mosiman, C., Poplawski, M., Saha, A., Bender, J., and Granderson, J. (2021). Metadata schemas and ontologies for building energy applications: A critical review and use case analysis. *Energies*, 14(7):2024.
- Rasmussen, M. H., Lefrançois, M., Schneider, G. F., and Pauwels, P. (2021). Bot: The building topology ontology of the w3c linked building data group. *Semantic Web*, 12(1):143–161.
- SEAI (2023). *Building Energy Rating Certificate Database* by SEAI. <https://ndber.seai.ie>.
- Shi, J., Pan, Z., Jiang, L., and Zhai, X. (2023). An ontology-based methodology to establish city information model of digital twin city by merging bim, gis and iot. *Advanced Engineering Informatics*, 57:102114.
- Terkaj, W., Schneider, G. F., and Pauwels, P. (2017). Reusing domain ontologies in linked building data: the case of building automation and control. In *8th International workshop on formal ontologies meet industry*, volume 2050.
- TUWien (2024a). *Building information ontology*. https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/building/1_10/gbBuildingOntology.owl. [Online; accessed 01-Feb-2024].
- TUWien (2024b). *Weather ontology*. <https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/WeatherOntology.owl>. [Online; accessed 01-Feb-2024].
- Van Dam, K. H. and Keirstead, J. (2010). Re-use of an ontology for modelling urban energy systems. In *Next generation infrastructure systems for eco-cities*, pages 1–6. IEEE.
- Wu, J., Orlandi, F., AlSkaif, T., O'Sullivan, D., and Dev, S. (2021). Ontology modeling for decentralized household energy systems. In *2021 SEST*, pages 1–6. IEEE.
- Wu, Z., Cheng, J. C., and Wang, Z. (2022). An ontology-based framework for building energy simulation in the operation phase. In *International Conference on Computing in Civil and Building Engineering*, pages 351–366. Springer.
- Zadeh, P. A., Wei, L., Dee, A., Pottinger, R., and Staub-French, S. (2019). Bim-citygml data integration for modern urban challenges. *Journal of Information Technology in Construction*, 24.
- Zhu, J., Chong, H.-Y., Zhao, H., Wu, J., Tan, Y., and Xu, H. (2022). The application of graph in bim/gis integration. *Buildings*, 12(12):2162.