A UTILITY-AWARE KNOWLEDGE GRAPH TO SUPPORT MULTI-SCALE RESIDENTIAL ENERGY MODELLING

Cathal Hoare[1], Sharon Coffee[2], Divyanshu Sood[2], Tiziana Margaria[1] and James O'Donnell[3]
[1] Department of Computer Science and Information Systems, University of Limerick, Ireland
[2] School of Material and Mechanical Engineering, UCD Energy Institute, University College Dublin, Ireland
[3] School of Engineering, University of Limerick, Ireland

Abstract

Building operations account for major global energy use, yet decarbonisation is hindered by fragmented data silos. This paper presents a utility-aware knowledge graph developed for multi-scale residential energy modelling. By leveraging Linked Data principles and the DDIM server, the framework integrates building-level performance data with national electricity distribution networks. Using Ireland as a case study, the methodology semantically associates 2.3 million residences with over 46,000 low-voltage transformers through a novel enrichment process. This scalable approach enables planners to perform complex queries for grid capacity analysis, bridging the gap between individual building physics and urban-scale energy simulation.

Introduction

Building operations accounted for 30% of global energy consumption in 2021 (Skillington et al., 2022; Bacheva and Raposo Grau, 2025). Ireland, like many countries, is on-track to miss decarbonisation goals, a potential significant cost to the exchequer (Environmental Protection Agency, 2025). Ireland is also suffering a housing crisis, with development of residential stock failing to meet demand (Social Justice Ireland, 2025). Infrastructure problems, such as insufficient electricity supply capacity, are a contributing factor to both of these issues (Irish Academy of Engineering, 2025). Data-driven decisions are now vital to address the issues of global warming, building comfort and stock availability (Skillington et al., 2022). By fusing various information sources, such as building data, energy infrastructure and population data, and applying Machine Learning (ML), planners can perform sophisticated analyses, to identify energy performance trends and optimize large-scale carbon reduction strategies (Ali et al., 2021a).

Access to accurate information is of fundamental importance to all stakeholders in this process, including planners and policy makers (Madnick et al., 2009). The information that must be managed is, however, complex and heterogeneous, both in content and organization (Bonilla et al., 2019). Information sources are dynamic, evolving, and multifaceted; information can be distributed across multiple sources providing related, yet heterogeneous collections of data that inform utility companies, planners and other stakeholders (Boddy et al., 2007). National and district level views of the same entities are managed by separate formats and systems, requiring expensive data reorganization in order to extract relevant information for planners and policy makers (Swan and Ugursal, 2009; Nacht et al., 2023). Together these issues describe rich, timely data sources whose utility is diminished by their disjointed and heterogeneous nature. This work presents an approach that seeks to overcome these difficulties and increases the value of data by integrating diverse sources into a homogeneous data source that can be extended to include new sources as they become available (Li et al., 2019).

Several advances have enabled solutions to these problems. Linked data provides a mechanism to interlink disparate datasources and interrogate them using semantic queries - that is queries that are of associative and contextual nature (Li et al., 2019). Linked data allows all relationships in a datastore to be explicitly defined. This facilitates the development of flexible queries, and admits reasoning about both data and their relationships (Ngomo et al., 2014). Linked data also permits the integration and exchange of data from a variety of sources and domains (Consoli et al., 2017). Digital twins for multi-scale energy modelling integrate Linked Data and ML to bridge the gap between building-level physics and urban-scale dynamics. Linked Data, utilizing RDF and ontologies, provides a semantic framework to unify heterogeneous data sources, such as IoT sensors, GIS, and weather APIs, ensuring interoperability across scales. Meanwhile, ML models accelerate simulations by acting as surrogate models for complex thermodynamic equations. By training on historical and real-time semantic data, these twins enable predictive analytics for grid stability and demand response. This synergy allows for a "glass-box" approach where data remains structured and searchable, while ML provides the computational power to optimize energy distribution in complex, multi-layered environments (Santos et al., 2025).

This research paper details the creation of a utility-aware knowledge graph designed to improve multi-scale energy modelling within the built environment. By utilizing Linked Data principles, the authors successfully integrate fragmented datasets involving building characteristics, population statistics, and electrical grid infrastructure. The study specifically demonstrates this approach in Ireland, connecting millions of residential properties to their nearest low-voltage transformers to create a unified view of energy demand. This framework addresses the problem of functional silos by allowing planners to perform sophisticated, data-driven analyses of decarbonisation strategies and grid capacity. Ultimately, the methodology provides a scalable solution for integrating heterogeneous information sources to support sustainable urban planning across different jurisdictions.

The paper is organized as follows. First, the background explores functional silos in the energy modelling domain and the application of Linked Data and standardized ontologies. The methodology then details the development of the utility-aware knowledge graph, focusing on the core context and the DDIM server used for managing data fusion. A central component describes the enrichment process that semantically links building assets to low-voltage electricity distribution networks. The results section examines the national-scale graph generated using Irish datasets, including GeoDirectory and ESB Networks. The study's capabilities are then validated through competency questions that demonstrate sophisticated reasoning. Finally, the paper discusses the methodology's global scalability before concluding with future research directions.

Background

Multi-scale residential energy-modelling is significantly hindered by functional silos, which are compartmentalised operating units isolated from their environment (Ahmad et al., 2018; Ali et al., 2021b). These silos result in horizontal fragmentation between project disciplines and vertical fragmentation between individual building and district-scale representations (Ali et al., 2024; Balaji et al., 2016). Data integration remains a vexing problem because building information is complex, dynamic, and distributed across multifaceted sources in incompatible formats (He et al., 2023). Planners frequently encounter data inconsistencies and a lack of standard building identifiers, leading to property discrepancies across datasets (Hoare et al., 2022; Pauwels et al., 2017). Access to high-quality energy data is further restricted by privacy and protection policies, which may exclude critical information or compromise its quality through anonymisation (Pritoni et al., 2021; Rasmussen et al., 2021). These hurdles necessitate tedious manual data reorganisation and collection efforts that are both time-consuming and costly (Sood et al., 2023; Tardioli et al., 2017). Consequently, these integration issues create a persistent gap between detailed building energy modelling and traditional urban planning practices (Wang et al., 2024; Wu et al., 2024).

Linked Data principles address the persistent challenge in multi-scale residential energy-modelling of functional silos by providing a mechanism to interlink disparate, heterogeneous data sources into a unified "Web of Data" (Pauwels et al., 2017; Hoare et al., 2022). By utilizing the Resource Description Framework (RDF) and unique Uniform Identifiers (URIs), this approach admits semantic queries of an associative and contextual nature, allowing all relationships in a datastore to be explicitly defined (Hoare et al., 2022; Bizer et al., 2009).

The sector has operationalized this through standardized ontologies:

- BOT provides a lightweight core for building topology (Rasmussen et al., 2021).
- Brick standardizes building automation metadata (Balaji et al., 2016).
- SAREF enables interoperability among smart appliances (Daniele et al., 2015).

A significant practical application is the use of federated SPARQL queries to integrate Ireland's national transmission grid data with residential housing stock archetypes (Hoare et al., 2022). This allows for utility-aware, multi-scale reasoning where data owners maintain local control while planners perform national-level scenario exploration for decarbonisation (Hoare et al., 2022).

Utility information in Linked Data is represented through specialized ontologies that bridge the gap between building-level assets and national infrastructure (Hoare et al., 2022). While standard building schemas like Brick often lack grid-specific concepts (He et al., 2023), extensions such as `SARGON` incorporate electrical grid automation by mapping CIM and IEC 61850 data models into RDF formats (Haghgoo et al., 2020).

Crucially, this information is structured using a hierarchical *common context* spine; individual buildings are semantically linked to geographical small areas and then to transmission nodes (Hoare et al., 2022; He et al., 2023).

Methodology

We now turn to describing the development of a utility aware knowledge graph (KG) to support multi-scale energy modelling. The KG is structured by defining a number of common contexts to create vertical hierarchies of relationships between represented entities and horizontal relationships to define equivalence between disparate representation of the same physical entities. While exemplified using the Irish context (and relevant datasets), the approach can be applied across jurisdictions where suitable datasets are available. Having described the common context, the process used to integrate building and electricity distribution network information is described. The section will conclude with a brief description of DDIM server, used to manage the data fusion and access process.

Core context

A graphical representation of the context model along with the organization of this project's data through the use of the ontology are shown in Figure 1 (A). The ontology is deliberately sparse to maintain clarity. An entity represents some physical entity or space in time. Each instance

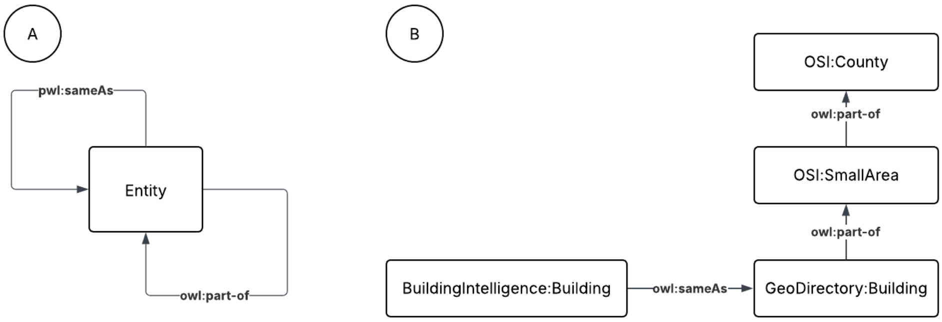

Figure 1: A. Context model and B. Relationships to data sets used in this project

of an entity is described by a class, for example, a building, region or nation. Together, these classes describe a set of layered contexts, each representing a different scale. Information sources are associated with some entity in each layer. This association with context allows sources to be intra-connected at a particular scale, and inter-connected between scales. Figure 1 (B) illustrates the use of the context model to place buildings in geographical hierarchies and relate building representations to each other across datasets. Equivalence is expressed using the *owl:sameAs* relationship; we use the relationship to denote that buildings in different datasets represent the same physical entity (Halpin et al., 2010). The use of *owl:sameAs* can give rise to semantic ambiguity where datasets' properties lists overlap. This issue is overcome by assigning a source authority ranking for datasets as they are added to the graph. This feature is provided through the DDIM's data upload forms. Hierarchies are created using the *owl:part-of* relationship. In Figure 1 building representations are in the geographical context of a small area which in turn is part of a county. In this way, various transitive relationships can be created to support aggregation and other analysis operations over geographical contexts.

Managing data fusion and access

The Dynamic District Information Model (DDIM) (Project, 2023a) is used to mine and manage relationships between the siloed data sources. For this project, the DDIM uses the core context to defined relationships between geographical areas and locations to provide a common context between data sources. This context is captured through a knowledge graph defiend in Neo4J (with NeoSemantic plugin). The server supports forms based no-code implementation of the graph, where forms are used to upload datasets and create owl:sameAs and owl:part-of relationships between columns in the uploaded data. the The server implementation consists of three key functional areas:

- Creation of 'Core Spines' (describe by the context model) using a form based process integrates uploaded information. These spines allows information to be associated with entities in various contexts, facilitating querying across these spaces;
- A series of interfaces to allow client software to query (through RESTful or SPARQL based interfaces);
- A DCAT compliant data catalogue;

The server is implemented using the Django Framework (Project, 2023b) with RESTful (DjangoRESTful, 2023) extensions. This framework also provides user management and security. Neo4J and NeoSemantic (Project, 2023a) is installed within the same domain. Queries are submitted through the Django server to this database.

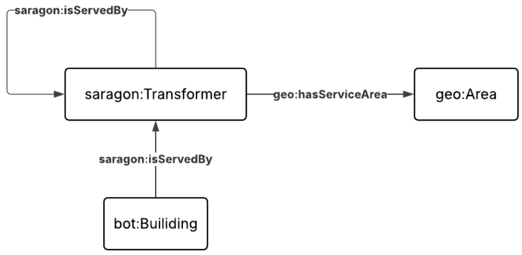

Figure 2: Additional graph structure added post enrichment.

Enriching KG with building/LV relationships

We now examine the steps taken to integrate building and electricity distribution network information. In this instance, the information used was provided by ESB Networks as part of their *Capacity Heatmap for Network Planning* tool (ESB Networks, 2026). This provides a wealth of information about the location and capacity of transformers on the Irish electricity distribution network. One purpose of the Capacity Heatmap is to guide customers who wish to make a new connection to the network. In this use, customers are advised to locate the transformer closest to the required connection; we employ the same heuristic in this work. Low Voltage (LV) transformers were associated with building data (from Geodirectory (GeoDirectory, 2022a)) and areas representing each transformer's serviced area were created using the following steps:

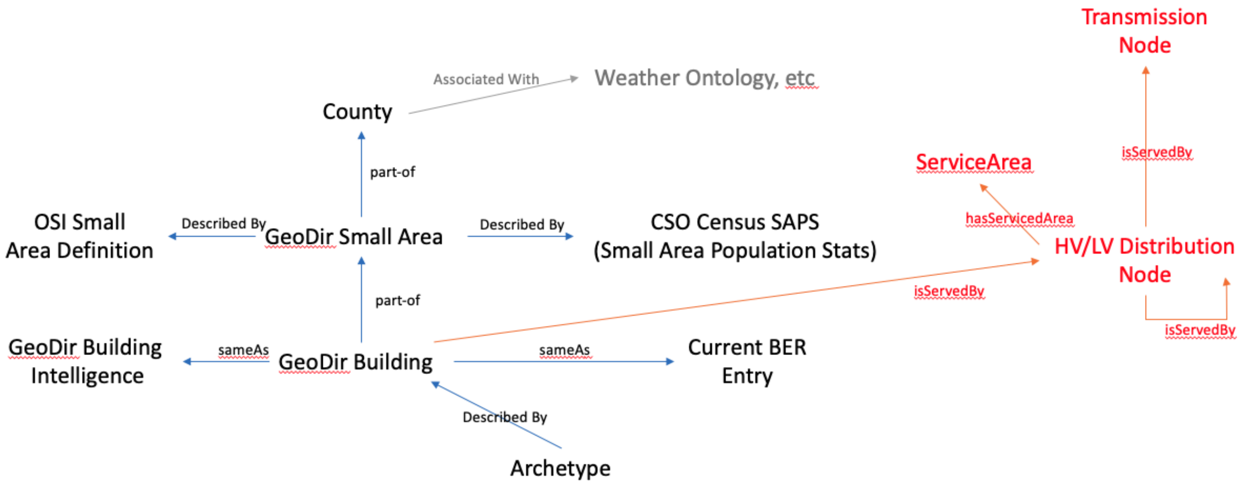

Figure 3: Representation of relationships between datasets in National Scale Residential Energy Graph. Elements added as part of the enrichment process are shown in red.

1. Entity definitions

Let B be the set of buildings:

$$B = \{b_i \mid i = 1, \ldots, n\}$$

where $b_i = (\mathbf{p}_i, \mathscr{C}_i)$, with $\mathbf{p}_i \in \mathbb{R}^2$ as coordinates and $\mathscr{C}_i$ as characteristics.
Let T be the set of transformers:

$$T = \{t_j \mid j = 1, \ldots, m\}$$

where $t_j = (n_j, \kappa_j, \mathbf{l}_j)$, with n_j as name, κ_j as capacity, and $\mathbf{l}_j \in \mathbb{R}^2$ as location.

2. Relationship assignment

For each building b_i, the closest transformer t^* is found via:

$$f(b_i) = \arg\min_{t_j \in T} d(\mathbf{p}_i, \mathbf{l}_j)$$

This establishes the relation $(b_i, \text{isServedBy}, f(b_i))$.

3. Geometric representation

The service area for transformer t_j is defined by the convex hull of the points S_j it supplies:

$$S_j = \{\mathbf{p}_i \mid f(b_i) = t_j\}$$

$$\text{CH}(S_j) = \left\{ \sum_{k=1}^{|S_j|} \alpha_k \mathbf{p}_k : \sum \alpha_k = 1, \alpha_k \geq 0 \right\}$$

4. Knowledge graph insertion

The graph $\mathscr{G}$ is updated such that:

$$\mathscr{G} \in \{(b_i, \text{isServedBy}, t_j), (t_j, \text{hasServiceArea}, \text{CH}(S_j))\}$$

These steps were implemented as a Python script which output two files, a text file containing Cypher queries to create a relationship between each building and the transformer that it is served by. The second output file is a GeoJSON (Internet Engineering Task Force (IETF), 2016) formatted definition of the convex hull defined areas that describes the area containing buildings serviced by a transformer. In the Irish context, 46,527 transformers were associated with over 2.3 million residences.

Figure 2 shows the structure of the developed graph extension when the LV/Building relationship process is processed. Transformers and their associated connections to buildings are described using the SmArt eneRGy dOmain oNtology (SARAGON) (Haghgoo et al., 2020) ontology. Transformers can be served by higher voltage transformers and in turn serve buildings. Each transfomer is also associated with a geographical area representing the convex hull described by the buildings it serves. This area can be used for visualisation and other purposes.

Results

This paper continues by examining the generated utility aware graph. The data sources and resulting graph structure will be described in the Irish context before examining the outputs Building/LV enrichment process. The convex hull outputs will be described before a number of competency questions will demonstrate the complexity of queries that the knowledge graph can inform. Finally, the potential to use this methodology in other jurisdictions will be discussed.

Structure of knowledge graph

The DDIM and core context were used to create a national scale residential energy graph for Ireland. The graph structure is used to inform energy modelling. Subsequently, the graph was enriched to include information about the Irish electricity distribution network using the steps described a-priori. Development of the graph followed the following steps:

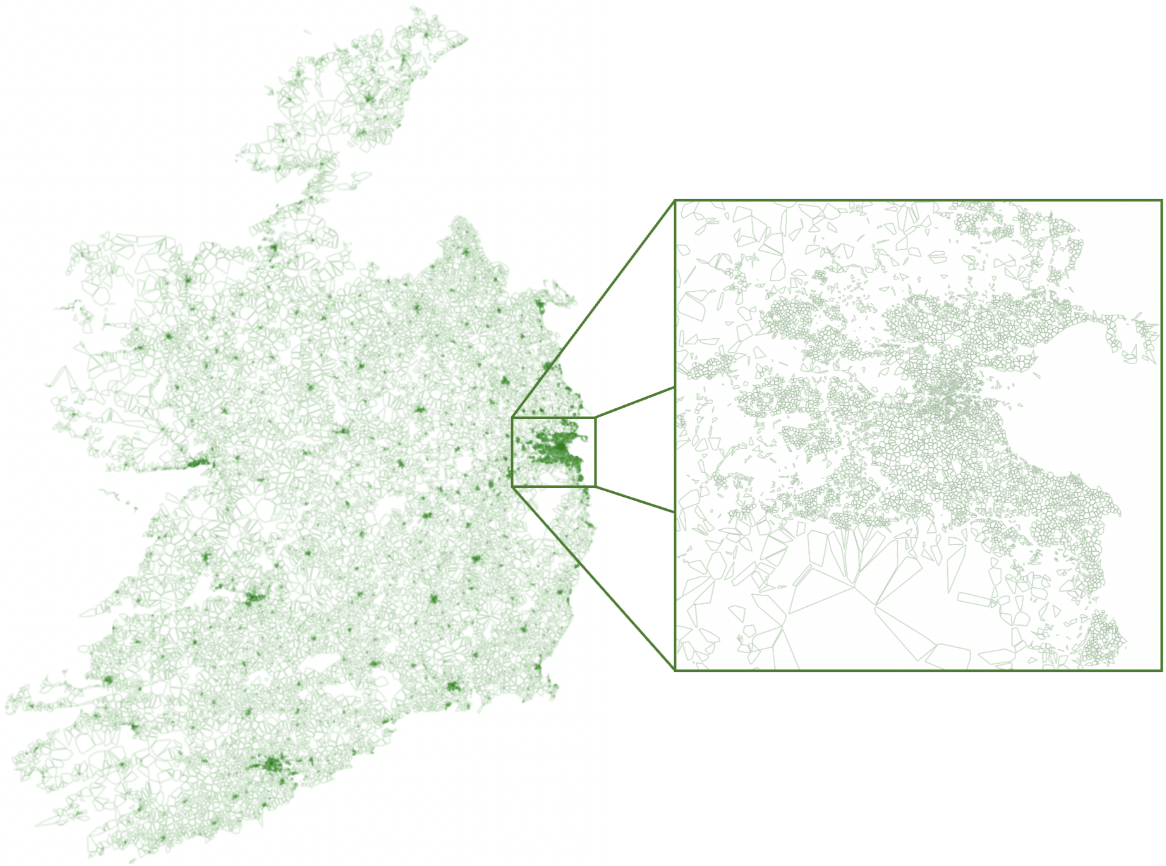

Figure 4: Map generated from presentation of derived convex hulls representing areas served by LV transformers. Clusters of smaller areas correspond to urban areas with higher densities of residential properties.

- Collection of relevant data sources.
- Identification of vertical and hierarchical relationships that could be exploited to support dataset fusion using the core context.
- Creation of the graph using the DDIM server.

The following datasets were identified:

- Geodirectory's *GeoAddress (GeoDirectory, 2022a)* and associated *BuildingIntelligence (GeoDirectory, 2022b)* – comprehensive list of approximately 2.3 million buildings, including location information and residential building properties. BuildingIntelligence includes energy modelling relevant information, for example, heating fuel.
- SEAI's *National BER Research Tool (SEAI, 2023)* - contains anonymised data used to calculate buildings' EPC certificate. The database contains data collected during building surveys. There are entries for approximately 60% of the Irish residential building stock.
- Irish Central Statistics Office (CSO) *Small Area Population Statistics (SAPS) (Office, 2022)* – 15 themes describe various statistics garnered from the CSO's census. These include socio-economic information about housing, commuting, motor cars and other energy modelling resources.

The Knowledge Graph contains other relevant information, including geographic area definitions from Ordinance Survey Ireland and developed building archtypes.

The DDIM was used to create the graph structure shown in Figure 3. Horizontal *same-as* relationships were identified between, building representations in the BER, GeoAddress and Building Intelligence datasets. Each building (represented by the GeoAddress entry) was associated with a building archtype. A vertical *core context* was created where each building was placed in the context of a small area using part-of relationships. Small areas were similarly placed in counties. These geographic contexts were chosen as they are widely used by various data providers. However, finer grained contexts including streets and townlands could also have been created.

The graph was subsequently enriched by creating associations between buildings and transformers (shown in red in Figure 3). Having followed the process described a priori. This added nodes for each transformer and a representation of the area they serve. Relationships between building representations and transformers were also created.

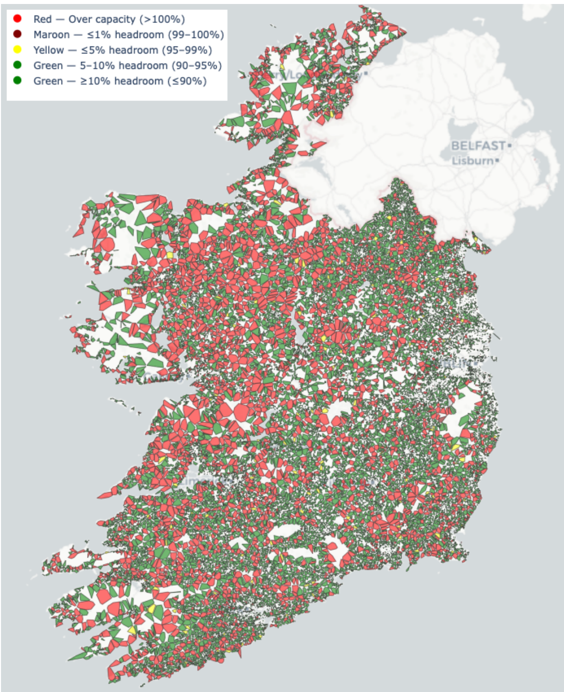

Figure 5: esults of multiscale energy modelling for scenarios including heatpump and EV adoption by residential households.

Visualising the relationships between LV transformers and buildings

Figure 4 presents a visualisation of the areas served by individual LV transformers at various scales. A distinct contast exists between urban areas characteristed by denser clusters of smaller areas and remote low-population areas (west coast and mountainous areas such as South of Dublin). This becomes more apparent when the map is zoomed to an urban areas such as Dublin (inset, Figure 4. As these areas are defined as GeoJSON polygons, colours can be assigned to each area depending on the form of analysis applied; for example, the visualisation can be used to show which LV transformers will experience capacity issues under various scenarios. The use of GeoJSON is also advantageous as web-based interfaces can be created dynamically zoom in and out across multiple-scales.

An example of this application is shown in Figure 5. In this isntance, building characteristics (fabric, occupancy and heating system) are aggregated into representative archetypes and converted to hourly electricity demand. Electrification scenarios such as heat-pump adoption and EV charging are applied as behavioural state changes within the housing stock. The resulting demand is spatially allocated to low-voltage substations, allowing feeder peak load and transformer overload to be calculated.

Competency questions for knowledge graph

Scoping and design of the KG was guided by developing numerous competency questions (Keet and Khan, 2024). These questions were also used to validate the capabilities of the graph. We now present two of these along with their results to demonstrate the ability of the KG to reason in sophisticated ways about knowledge relevant to multi-scale energy modelling in Ireland.

CQ1: Average residential buildings per LV node

The following Cypher query was used to report on statistics related to this query:

```
MATCH (b:Building)-[:NEAREST_LV_STATION]->(
lv:LVStation)

WITH lv, count(b) AS buildingCount

RETURN avg(buildingCount) AS
avgBuildingsPerLVNode,
min(buildingCount) AS minBuildings,
max(buildingCount) AS maxBuildings,
count(lv) AS totalLVNodes
```

The average number of connections was found to be 47.0 buildings per LV node with a standard deviation of 44.5. In total 42,235 LV nodes were found to supply at least one residential building. This can be explained by some LVs supplying commercial buildings and the use of the nearest LV to a building heuristic.

CQ2: Aggregate connectivity summary

Provide a summary table that describes overall connectivity between LV nodes and residential buildings.

```
MATCH (lv:LVStation)
OPTIONAL MATCH (lv)<-[:NEAREST_LV_STATION]-
                        (b:Building)
WITH lv, count(DISTINCT b) AS buildings

OPTIONAL MATCH (lv)-[:SERVES_SMALL_AREA]->
                        (sa:SmallArea)
WITH lv, buildings, count(DISTINCT sa) AS
                            smallAreas

OPTIONAL MATCH (lv)-[:HAS_ARCHETYPE_COUNTS]
->(ac:ArchetypeCounts) WITH lv, buildings,
smallAreas, count(DISTINCT ac) AS
                         archetypeNodes
WITH buildings + smallAreas +
     archetypeNodes AS totalConnections

RETURN avg(totalConnections) AS
avgConnectivity, min(totalConnections)
AS minConnectivity, max(totalConnections)
AS maxConnectivity, percentileCont(
totalConnections, 0.5) AS
medianConnectivity
```

In summary, overall connectivity is described thus: average 53.9, median 41.0, range 0–619. The sophistication of the query and its ability to bridge disparate siloed information sources is demonstrated.

Applying this methodology in other jurisdictions

The methodology is designed for global scalability. Access to appropriate datasets represents the greatest challenge. For example, in the United Kingdom, access to GeoPlace's National Address Gazetteer would yield a dataset of addresses equivalent to GeoDirectory, while EPC data can be accessed through the Department for Levelling Up, Housing & Communities' Energy Performance of Buildings Data datasets. By mapping these and other datasets to common contexts, the no-code form based approach used by the DDIM allows domain experts to create utility-aware knowledge graphs in any jurisdiction. Larger datascales at national level are supported by the ability of the DDIM to operate as a federated node in distributed deployment. Using building archetypes further enables national-scale energy simulations for diverse decarbonisation planning.

Conclusions

This paper addresses the fragmentation of AECOO data silos by developing a utility-aware knowledge graph (KG) for multi-scale residential energy modelling. By leveraging Linked Data principles and the Dynamic District Information Model (DDIM), the research successfully integrated building-level characteristics with national electricity infrastructure. The methodology was validated in the Irish context, semantically linking 2.3 million buildings to 46,527 Low Voltage (LV) transformers through a novel automated enrichment process. Results from competency questions demonstrated the KG's ability to perform sophisticated reasoning, revealing an average of 47.0 residential buildings served per LV node. This framework effectively bridges building physics with urban-scale dynamics, offering a globally scalable approach for planners to optimize grid capacity and decarbonisation strategies. Future work will focus on integrating real-time sensor data and machine learning to enhance predictive analytics for grid stability

Acknowledgments

This publication has emanated from research supported partly by a grant from the Sustainable Energy Authority of Ireland (SEAI) under the Research, Development and Demonstration Funding Programme Grant No. 22/RDD/1300. 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 SEAI. Anthropic Claude used to improve the grammar and clarity of some text here-in; all content are the authors' responsibility.

References

Ahmad, T., Chen, H., Guo, Y., and Wang, J. (2018). A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: A review. Energy and Buildings, 165:301–320.

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.

Ali, U., Shamsi, M. H., Bohacek, M., Purcell, K., Hoare, C., and O'Donnell, J. (2021a). Gis-based multi-scale residential building energy modeling using a data-driven approach. In Proc. IBPSA 2021, volume 17 of Building Simulation, pages 1115–1122, Bruges, Belgium. IBPSA.

Ali, U., Shamsi, M. H., Hoare, C., Mangina, E., and O'Donnell, J. (2021b). Review of urban building energy modeling (ubem) approaches, methods and tools using qualitative and quantitative analysis. Energy and Buildings, 246:111073.

Bacheva, T. S. and Raposo Grau, J. F. (2025). Embodied impacts in buildings: A systematic review of life cycle gaps and sectoral integration strategies. Buildings, 15(10).

Balaji, B., Bhattacharya, A., Fierro, G., et al. (2016). Brick: Towards a unified metadata schema for buildings. In Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments (BuildSys '16), pages 41–50.

Bizer, C., Heath, T., and Berners-Lee, T. (2009). Linked data - the story so far. International Journal on Semantic Web and Information Systems (IJSWIS), 5(3):1–22.

Boddy, S., Rezgui, Y., Cooper, G., and Wetherill, M. (2007). Computer integrated construction: A review and proposals for future direction. Adv. Eng. Softw., 38(10):677–687.

Bonilla, J., Perez, J., Pinto-Seppa, I., and Zarli, A. (2019). Articial intelligence - research and innovation needs of manufacturing, energy intensive industries, bio-based industries and construction. Technical report, The European Construction Technology Platform - and the Energy Efficient Building.

Consoli, S., Presutti, V., Recupero], D. R., Nuzzolese, A. G., Peroni, S., Mongiovi', M., and Gangemi, A. (2017). Producing linked data for smart cities: The case of catania. Big Data Research, 7:1 – 15.

Daniele, L., den Hartog, F., and Roes, J. (2015). Created in close interaction with the industry: The smart appliances reference (saref) ontology. In Proceedings of the 7th International Workshop on Formal Ontologies Meet Industry (FOMI 2015), pages 100–112. Springer.

DjangoRESTful (2023). Django rest framework. `https://www.django-rest-framework.org`. [Online; accessed 27-Feb-2023].

Environmental Protection Agency (2025). EPA projections show Ireland off track for 2030 climate targets. `https://www.epa.ie/news-releases/news-releases-2025/epa-projections-show-ireland-off-track-for-2030-clima` php. Accessed: 2026-02-26.

ESB Networks (2026). Network Capacity Heatmap. Accessed: 2026-02-23.

GeoDirectory (2022a). Geoaddress smart data. https://www.geodirectory.ie/products-services/geoaddress-smartdata. [Online; accessed 27-Feb-2023].

GeoDirectory (2022b). Geobuilding intel. https://www.geodirectory.ie/products-services/geobuilding-intel. [Online; accessed 27-Feb-2023].

Haghgoo, M., Sychev, I., Monti, A., and Fitzek, F. H. (2020). Sargon – smart energy domain ontology. IET Smart Cities, 2(4):191–198.

Halpin, H., Hayes, P. J., McCusker, J. P., McGuinness, D. L., and Thompson, H. S. (2010). When owl:sameas isn't the same: An analysis of identity in linked data. In Patel-Schneider, P. F., Pan, Y., Hitzler, P., Mika, P., Zhang, L., Pan, J. Z., Horrocks, I., and Glimm, B., editors, The Semantic Web – ISWC 2010, pages 305–320, Berlin, Heidelberg. Springer Berlin Heidelberg.

He, F., Wang, D., and Sun, Y. (2023). Ontology integration for building systems and energy storage systems. In Proceedings of the 10th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (BuildSys '23), pages 445–448.

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.

Internet Engineering Task Force (IETF) (2016). RFC 7946: GeoJSON Specification. Accessed: 2026-02-23.

Irish Academy of Engineering (2025). Rebalancing ireland's energy policy. Technical report, Irish Academy of Engineering, Dublin, Ireland.

Keet, C. and Khan, Z. (2024). Discerning and characterising types of competency questions for ontologies.

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.

Madnick, S. E., Wang, R. Y., Lee, Y. W., and Zhu, H. (2009). Overview and framework for data and information quality research. J. Data and Information Quality, 1(1).

Nacht, T., Pratter, R., Ganglbauer, J., Schibline, A., Aguayo, A., Fragkos, P., and Zisarou, E. (2023). Modeling approaches for residential energy consumption: A literature review. Climate, 11(9).

Ngomo, A.-C. N., Auer, S., Lehmann, J., and Zaveri, A. (2014). Introduction to Linked Data and Its Lifecycle on the Web, pages 1–99. Springer International Publishing, Cham.

Office, C. S. (2022). Small Area Population Statistics. Accessed: 2026-02-23.

Pauwels, P., Zhang, S., and Lee, Y.-C. (2017). Semantic web technologies in aec industry: A literature overview. Automation in Construction, 73:145–165.

Pritoni, M., Paine, D., Fierro, G., et al. (2021). Metadata schemas and ontologies for building energy applications: A critical review and use case analysis. Energies, 14(7):2024.

Project, D. (2023a). Django project. https://www.djangoproject.com. [Online; accessed 27-Feb-2023].

Project, D. (2023b). Django project. https://www.djangoproject.com. [Online; accessed 27-Feb-2023].

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(3):437–461.

Santos, R., Costa, Á., Silva, J., et al. (2025). Integrating machine learning and digital twins for enhanced smart building operation and energy management: A systematic review. Urban Science, 9(6):202.

SEAI (2023). Ber research too. https://ndber.seai.ie/BERResearchTool/ber/search.aspx. [Online; accessed 27-Feb-2023].

Skillington, K., Crawford, R. H., Warren-Myers, G., and Davidson, K. (2022). A review of existing policy for reducing embodied energy and greenhouse gas emissions of buildings. Energy Policy, 168:112920.

Social Justice Ireland (2025). National housing plan 2025-2030: Submission to department of housing, local government and heritage. Technical report, Social Justice Ireland, Dublin, Ireland.

Sood, D., Alhindawi, I., Ali, U., McGrath, J. A., Byrne, M. A., Finn, D., and O'Donnell, J. (2023). Simulation-based evaluation of occupancy on energy consumption of multi-scale residential building archetypes. Journal of Building Engineering, 75:106872.

Swan, L. G. and Ugursal, V. I. (2009). Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. Renewable and sustainable energy reviews, 13(8):1819–1835.

Tardioli, G., Kerrigan, R., Oates, M., O'Donnell, J., and Finn, D. (2017). A data-driven modelling approach for large scale demand profiling of residential buildings. In Proceedings of BS2017: 15th Conference of International Building Performance Simulation Association, San Francisco, CA, USA.

Wang, M., Lilis, G. N., Mavrokapnidis, D., et al. (2024). A knowledge graph-based framework to automate the generation of building energy models using geometric relation checking and hvac topology establishment. Energy and Buildings, 325:115035.

Wu, Z., Cheng, J. C. P., and Wang, Z. (2024). An ontology-based framework for building energy simulation in the operation phase. In Advances in Information Technology in Civil and Building Engineering: Proceedings of ICCCBE 2022, volume 357 of Lecture Notes in Civil Engineering, pages 351–366. Springer Nature Switzerland.