Developing an ontology-based tool for relating risks to the energy performance gap in buildings

Article

Accepted Version

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

1   **Developing an ontology-based tool for relating risks to the energy performance gap in**
2   **buildings**
3
4   **Abstract**

5   **Purpose-** Despite extensive research on the underlying reasons for the energy performance gap
6   in buildings, there is a critical need for stakeholders to standardize and facilitate the use of this
7   knowledge and support its broader application by machines. Our research addresses this gap by
8   developing both an ontology and a tool to utilize risk information regarding the performance
9   gap in buildings.

10   **Design/methodology/approach-** Research into this topic began with the creation of an energy
11   performance gap-risk ontology for new and existing buildings using the METHONTOLOGY
12   method. This comprised a comprehensive literature review and semi-structured interviews with
13   ten experts concerning six buildings, in order to develop taxonomies and define risk factor
14   interactions. It was followed by a three-stage validation using a mixed-method research
15   methodology. Steps included comparing the ontology with a similar empirical study, gathering
16   expert opinions via interviews and ratings assessments, and finally, interviewing an experienced
17   professional to ascertain whether there were any concepts not covered by the ontology. The
18   taxonomies were modeled in Protégé 5.5, and using the ontology, a spreadsheet tool was
19   developed using Microsoft Visual Basic for Applications in Excel.

20   **Findings-** The ontology identified 36 primary risk factors, and a total of 95 when including
21   additional risks linked to certain factors. Factors such as professional liability insurance,
22   stakeholder motivation, effective communication, experience, training, integrated design,
23   simplicity of detailing, building systems or design, and project commissioning can help manage
24   the performance gap in buildings. The tool developed serves as a decision-support system,
25   offering features like project risk checklists to assist stakeholders in addressing the performance
26   gap.

27   **Quality/value-**This study is the first to develop an energy performance gap-risk ontology and
28   a tool to help project stakeholders collect, store, and share building risk information.

29   **Keywords**
30   Energy performance gap, ontology, spreadsheet tool, project risks, risk identification
31   **Paper type** Research paper

1.Introduction

Buildings are responsible for significant energy consumption and energy-related greenhouse gas emissions (Alam *et al*., 2017). Therefore, it is critical to plan the right policies to improve the energy efficiency of new and existing building stock (Burman *et al*., 2014). To address this problem, governments have upgraded energy and construction standards in buildings and energy performance assessment tools worldwide. These efforts have led to the emergence of a series of low-carbon and low-energy buildings, both newly built and retrofitted (Gupta *et al*., 2020).

Nevertheless, energy estimates at the design stage often differ from actual operational use, and this difference is known as the energy performance gap (Godefroy, 2022). The magnitude of the energy performance gap (EPG) varies widely (Shi *et al*., 2019). In reviewed publications, Mahdavi and Berger (2024) found a median EPG of +30% in residential and +14% in non-residential buildings, while Calì *et al.* (2016) reported that the EPG can be up to 287%.

This phenomenon impacts various aspects of the building industry, including governmental sustainability targets (Ortiz *et al*., 2020), design, economic, technological, well-being, and health benefits (Shrubsole *et al*., 2019). It also affects the credibility of industry professionals, such as policymakers, engineers, and designers (Wang *et al*., 2023). Additionally, energy performance risk has financial implications for energy service companies, which typically guarantee project savings through energy performance contracting (Doylend, 2015).

The EPG of buildings, including green buildings, has been extensively studied for over two decades (Shi *et al*., 2019), with significant efforts being made to identify its causes (Pomponi and Moncaster, 2018) and propose strategies to bridge the gap. However, current research focuses on the technical aspects of building energy performance to reduce EPG, frequently overlooking important social and organizational factors (Zheng *et al*., 2024).

Furthermore, some authors have identified risks contributing to the gap. Risk is characterized as uncertain events impacting project goals (Siraj and Fayek, 2019) and performance (Jayasudha and Vidivelli, 2016). Significant uncertainty persists both throughout the building's life cycle and when replicating actual conditions in energy simulations (Garwood, 2019). Therefore, reducing uncertainties and implementing risk management strategies early in construction increases the likelihood of achieving the project goals (Yousri *et al*., 2023) and effectively mitigates the energy performance gap (Frei *et al*., 2017).

However, relatively few studies examine the EPG issue from a risk perspective (Doylend, 2015; Alam *et al*., 2017; and Topouzi *et al*., 2019). Furthermore, while these studies provide valuable

65 insights into risk factors and their classification, they lack the comprehensive overview

66 necessary to account for the varied risks across different contexts since they focus on one

67 country, and one case study. Additionally, the findings of these studies often overlap with

68 previous research identifying the causes of EPG and exploring it through risk management

69 literature. These studies categorize risks into different classes and this redundancy in

70 terminology and classification hinders the effective communication and practical application of

71 the accumulated knowledge and expertise in current practice to reduce the gap in buildings.

72 Therefore, standardization in the EPG domain, particularly from a risk perspective, is necessary

73 for effective energy performance gap mitigation.

74 Developing an ontology is often considered the first step towards harmonizing domain

75 knowledge across various information systems (Jiang *et al.*, 2023). Ontologies provide benefits

76 such as semantic modeling, reusability, and the extensibility of information (Schachinger and

77 Kastner, 2017; Han *et al.*, 2015). However, despite the existence of several ontologies in

78 building energy efficiency (Tah and Abanda, 2011; Corry *et al.*, 2015; Zhou and El-Gohary,

79 2017), a gap remains in the ontological representation linking risks to the performance gap and

80 specifying interrelationships between risk factors across multiple building projects involving

81 different building uses. Moreover, the construction sector needs to work on capturing, storing,

82 sharing, and re-using knowledge due to a lack of mechanisms and processes that encourage the

83 necessary social interaction to shape and formalize it (Shelbourn *et al.*, 2006). Therefore, an

84 environment is needed that can not only standardize these processes in a structured manner, but

85 also serve as a guideline, and transfer risk knowledge to future projects.

86 Given these research gaps, the primary aim of this study is to develop an ontology to relate risks

87 to EPG. The objectives of the paper are to:

88 − establish a common vocabulary to eliminate heterogeneity when identifying EPG risks in

89   buildings;

90 − classify risk factors and define their interrelations;

91 − develop a tool to assist project stakeholders in gathering, storing, and sharing the risk

92   information of energy-efficient building projects.

93 Our research contributes to the existing body of knowledge by developing a comprehensive

94 ontology that synthesizes empirical and theoretical knowledge across different building types,

95 certification systems, and contexts. The ontology facilitates knowledge dissemination among

96 project stakeholders and ensures semantic interoperability. By leveraging the ontology into a

97 risk management tool, the research supports the systematic collection of data from buildings

98 and the mitigation of EPG, and contributes to the United Nations' sustainable development

goals (SDG). The first section of this paper introduces the study. The second section provides background information, focusing both on the reasons for and risks surrounding the gap and on previous ontology studies. The third section details the research methodology, while the fourth section presents research findings on the ontology and the tool developed. The fifth section offers a discussion, and the final section covers conclusions, research limitations, and future work.

2. Background

2.1. Causes of the energy performance gap

A widely accepted definition describes EPG as the difference between calculated (or simulated) and measured energy use (Bai *et al*., 2024), arising from concurrent factors present throughout a building's life cycle (Hahn *et al*., 2020). Researchers identified EPG factors through various methods, including literature reviews (Van Dronkelaar *et al*., 2016), surveys with facility managers (Liang *et al*., 2019), and detailed analyses of project documentation, thermography, co-heating tests, interviews, occupant surveys, and walkthroughs (Gupta *et al*., 2013).

In the design phase, EPG is influenced by limitations in modeling programs and methods (Menezes *et al*., 2012), misuse of tools (Kampelis *et al*., 2017), unrealistic behavioral assumptions (Gram-Hanssen and Georg, 2018), design complexity, early design choices, and human errors (Godefroy, 2022). Wang *et al*. (2023) highlight the lack of actual data on existing buildings and the disregarding of thermal bridges and insulation gaps during energy modeling. Factors such as post-design changes and construction quality can cause EPG in the construction phase, while unfinished activities and poor-quality handovers contribute to EPG at the commissioning and handover stage (Godefroy, 2022). During operation, occupant-driven factors predominantly cause EPG (Mahdavi & Berger, 2024), including higher operating temperatures, increased air change rates, and discrepancies in plug-loads, lighting usage, and internal heat loads. For this reason, the knowledge and skills of the occupants and energy managers are crucial (Zou *et al*., 2018). Further factors leading to EPG include poor practices, faulty equipment, measurement system limitations, operational instability, maintenance, and facility management issues (Godefroy, 2022).

In addition to the root causes of the gap, strategies for closing it are among the most widely studied areas in current research. Most researchers and practitioners consider technical methods, such as data collection and simulation processes, to be among the best ways to reduce the gap (Zheng *et al*.,2024), as well as transparency in energy performance data reporting and benchmarking (Danish & Senjyu, 2023). However, resolving the EPG also requires soft methods, such as effective communication and management among building stakeholders, and

133 mandatory regulatory strategies (Zheng *et al*., 2024). Therefore, effective stakeholder
134 engagement and collaboration (Madhusanka *et al*., 2022), along with strategies such as designer
135 competence, early involvement of key participants, and an integrated project delivery model,
136 are also critical to bridging the gap (Moradi *et al*., 2024).

137 *2.2. Risks influencing the gap*

138 Risk is often described in terms of uncertain events and their influence on project goals (Siraj
139 and Fayek, 2019). Therefore, early-stage risk identification helps ensure that stakeholders and
140 clients achieve their project goals (Yousri *et al*., 2023). The ISO 31000:2018 standard
141 emphasizes risk assessment—comprising identification, analysis, and evaluation—as central to
142 risk management.

143 Risk assessment models in green building projects are less comprehensive than in general risk
144 literature (Nguyen and Macchion, 2023). Mills *et al*. (2006) identified five classes of energy-
145 efficient project risks: measurement and verification, economic, operational, technological, and
146 contextual. Qin *et al*. (2016) examined certification, managerial, quality/technological,
147 financial/cost, political, and social risks in the green building life cycle in China, emphasizing
148 their probability and impact. Yang *et al.* (2016) showed that the critical risks for and
149 stakeholders of green buildings differ between countries (Australia and China).

150 The effective mitigation of EPG requires a well-structured, integrated performance and risk
151 management process (Frei *et al*., 2017). However, studies focusing on risks causing EPG are
152 limited. Doylend (2015) categorized energy performance risks into four groups: design and
153 engineering, management and process, external constraints, and operation and maintenance,
154 while Alam *et al*. (2017) categorized risks into six classes: design input, client-related issues,
155 procurement, construction management, material and equipment, and knowledge and skills.
156 Furthermore, Topouzi *et al*. (2019) identified three main risks: communication, sequence, and
157 assessment, comparing their likelihood in five retrofit approaches, and Thompson *et al*. (2022)
158 identified twenty-two risk factors in an analysis of 49 non-residential buildings.

159 *2.3. An overview of ontology studies*

160 Ontologies, sometimes described as vocabularies, contain a formalized representation of
161 knowledge for a particular domain in the information science field (Pritoni *et al*., 2021). A
162 hierarchy of concepts illustrating entity types, relations among concepts, restrictions on
163 relations, and instances are significant parts of ontologies (Schachinger and Kastner, 2017).
164 Ontologies facilitate knowledge exchange between domains and link shared knowledge,
165 offering advantages like semantic modeling (Schachinger and Kastner, 2017), information
166 reusability, extensibility, and interoperability (Han *et al*., 2015). They are useful in the research

167 areas of artificial intelligence, system integration, the semantic web, and problem-solving
168 methods (Tserng *et al.*, 2009).

169 Ontology development typically follows an iterative process with various modeling methods
170 (Schachinger and Kastner, 2017). Ontology building uses a customized procedure with no
171 universal method. Among the most common methods used in the construction industry are
172 METHONTOLOGY, SKEM, Uschold & Gruninger's (1996) approach, and NeOn, Grüninger,
173 and Fox's (1995) approach (Zhao *et al.*, 2016). Iqbal *et al.* (2013) conducted a comprehensive
174 review of fifteen ontology engineering methodologies and concluded that, while none of the
175 methodologies are fully mature, METHONTOLOGY stands out by providing detailed
176 descriptions of the techniques and activities employed.

177 Ontologies related to building energy efficiency serve multiple purposes. Researchers have
178 developed ontologies for selecting photovoltaic systems (Tah and Abanda, 2011), extracting
179 energy requirements from energy conservation codes (Zhou and El-Gohary, 2017), identifying
180 occupants' behavioral adaptation mechanisms (Hong *et al.*, 2015), and representing interactions
181 between smart grids and building energy management systems (Schachinger and Kastner,
182 2017). Other focuses include thermal comfort and energy efficiency (Esnaola-Gonzalez *et al.*,
183 2021) and performance assessment via a semantic web-based method (Corry *et al.*, 2015).

184 *2.4. Research contribution*

185 A comprehensive literature review on EPG research revealed the following critical limitations
186 in existing studies:

187 – Existing research predominantly focuses on the technical aspects of building energy
188   performance to mitigate EPG, often neglecting crucial social and organizational factors.

189 – Performance gap studies can be categorized into two groups: those with a risk management
190   perspective and those without. Despite using different terms like cause, reason, and risk, the
191   findings overlap significantly between these groups.

192 – Most studies in the risk management literature use a structured approach with risk
193   classification, something often lacking in EPG studies. Additionally, existing literature on
194   risk identification typically categorizes risks into different classes. The development and
195   application of classifications enhance communication efficiency by revealing patterns and
196   providing a comprehensive overview through the visualization of clusters, densities, and
197   gaps (Kwaśnik, 2019). However, inconsistent terminology and classification between
198   studies complicate the use of previous research insights.

199   –   Existing literature struggles to establish causal relationships between risk factors.

200   Nevertheless, it is essential to consider risk paths, both to prevent significant risks from

201   being disregarded (Alam *et al*., 2017) and to enhance risk mitigation.

202   –   Additionally, earlier studies on risks affecting building energy performance have been

203   constrained by focusing only on the UK construction sector, renovation methods, literature

204   reviews, and a single case study. However, previous researchers noted that risks affecting

205   building performance vary from one building to another (De Wilde, 2014), and critical risks

206   differ between different stakeholders and countries (Yang *et al*., 2016).

207   –   Current ontologies address the technical aspects of building energy performance; however,

208   no domain ontology systematically categorizes and defines the relationships between key

209   risks in EPG.

210   –   This study addresses current research limitations by developing an ontology that considers

211   various building types, sustainability standards, and country conditions to provide a

212   comprehensive view of risks affecting EPG. The ontology will standardize risk

213   terminology, classify risks systematically, and establish causal relationships between the

214   risks. Through semi-structured interviews considering the life-cycle stages of different

215   buildings, the study will explore not only technical but also social and organizational factors

216   causing EPG. Later, a tool will be developed to integrate risk management into the project

217   life cycle to reduce the gap in buildings. In this study, risks are defined as uncertain events

218   or situations that can impact building performance either negatively, positively, or both.

219 **3. Research steps and methods**

220 The study includes two main parts: (1) a five-step process for ontology development and (2)

221 the development of a tool based on the ontology. It proposes an ontology rather than a model

222 or conceptual framework, as ontologies represent knowledge, facilitate interoperability, and

223 allow semantic modeling. Although a conceptual framework outlines the current state of

224 knowledge, it is finalized before the study and is rarely modified once data collection begins

225 (Varpio *et al*., 2020).

226 Figure 1 illustrates the research steps employed in the study. The ontology was created using

227 the METHONTOLOGY method, as referenced by Zhou *et al*. (2016) and Guyo *et al*. (2023).

228 METHONTOLOGY is well-structured (Fernandez *et al*.,1997), comprehensive, and one of the

229 most frequently used ontology engineering methodologies (Abanda *et al*., 2017). It enables the

230 creation of an ontology from scratch (Abanda *et al*., 2017; Khalid *et al*., 2023), while also

231 permitting the reuse of existing ontologies. Due to the evolving prototype life cycle of this

232   methodology, ontology development is a continuous process, allowing updates at any phase
233   (Khalid *et al*., 2023). The ontology can be employed to create various tools suited to specific
234   requirements. This article provides an illustrative example. Following the ontology
235   development steps, a practical Excel-based tool, EPG-RISK, was created within a spreadsheet
236   environment to help project stakeholders collect, store, and share the risk information of
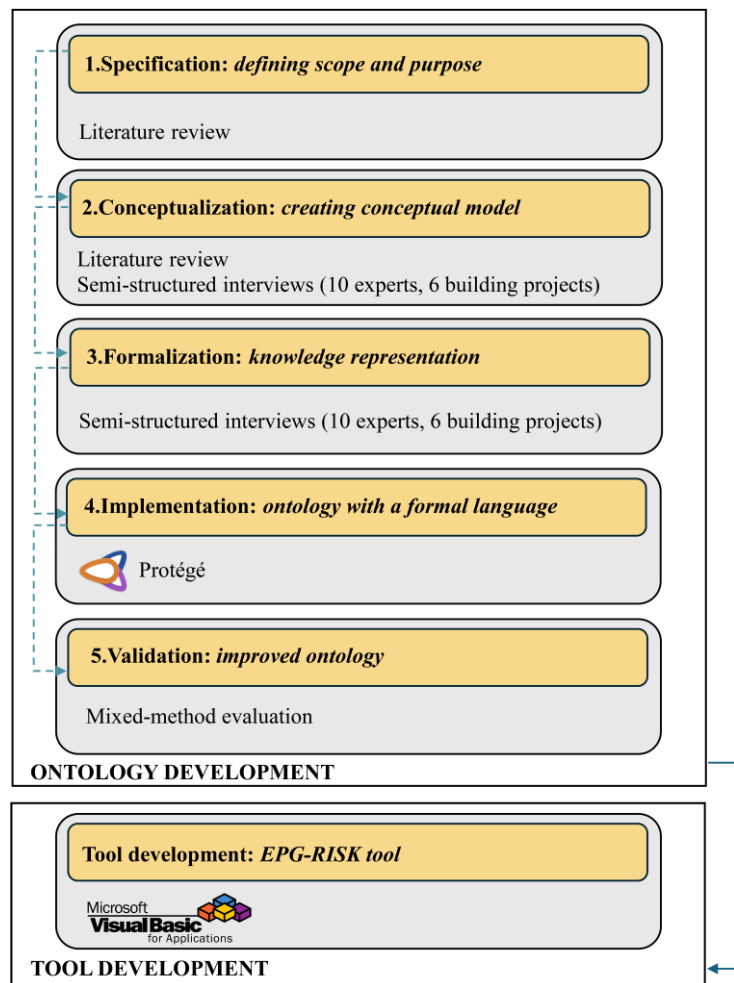237   projects.

Figure 1. Research steps (Source: Authors own work)

3.1 Ontology development stage

The ontology development process consists of five main steps: specification, conceptualization, formalization, implementation, and validation. The following sections explain each step in detail.

3.1.1 Specification

At a minimum, the specification step should provide the ontology's purpose, level of formality, and scope (Fernandez *et al*., 1997). This ontology aims to explain the energy performance gap in buildings by utilizing project risks. The ontology can then be used by (i) project managers, energy consultants, engineers, and energy service companies involved in developing a specific energy-efficient building project and assessing project risks, or (ii) experts who want to predict the risk of an energy performance gap in a project. Professionals can use the ontology to describe risks influencing EPG in a semi-formal language, considering the design, construction, and operational phases. Additionally, it helps identify relationships between various risk factors.

3.1.2 Conceptualization

The conceptualization process aims to uncover knowledge related to risks contributing to EPG in buildings. Conceptualization, a challenging aspect in ontology design, requires a subjective representation of the world and an understanding of how individuals perceive and categorize their environment (Fidan *et al*., 2011).

This step involved the identification of risks through an extensive review of the existing literature and semi-structured interviews concerning six building projects. Semi-structured interviews are frequently used to understand the 'what' and 'how', with a particular emphasis on the 'why'. Additionally, they help us understand the context and analyze relationships between variables (Saunders *et al*., 2019). Several researchers have employed semi-structured interviews (Moradi *et al*., 2024; Alencastro *et al*., 2024; Yousri *et al*., 2023), which was also the preferred method in this study as the aim was to understand the contextual factors for risk and EPG, particularly interrelations.

Initially, critical parameters, such as modeling, software, calculation methodology (De Wilde, 2014; Doylend, 2015; Calì *et al*., 2016), simulation inputs (De Wilde, 2014), and design problems (De Wilde, 2014; Doylend, 2015), were identified via a literature review. Twenty journal articles on EPG in buildings were reviewed, and the most common concepts collected. Later, semi-structured interviews were conducted with domain experts to explore factors affecting risk and EPG, understand their relationships, and develop a conceptual model. One criticism of semi-structured interviews is that the data collected may be perceived as "subjective and imprecise." However, conducting multiple meetings and interviews with the same respondents can enhance data quality and build trust (Albaret and Deas, 2023). Our study addressed these concerns by conducting two rounds of semi-structured interviews. The interviews were held between December 2020 and May 2021, either online or in person, each lasting 60 to 90 minutes. In the first round, interviewees were asked to describe the project phases of an energy-efficient building they had worked on, explaining problems or challenges that might result in an EPG, and stating whether these issues were resolved or led to further problems. In the second round, the identified risk factors and relationships were presented to the interviewees to determine their agreement, gather their feedback, and request suggestions for revisions.

The building project selection process was strategically designed to capture diverse perspectives on EPG in buildings applying the principles of sustainable design, both with and without certification. Projects in Turkey and Germany were selected to provide a comprehensive contextual lens. It is hypothesized that Turkey, offering the perspective of an

288   emerging market in green buildings, and Germany, as a pioneer, particularly in Passive House

289   certification, can both be representative and reflect different but complementary perspectives.

290   The projects that are discussed during the semi-structured interviews included one educational,

291   two residential, and three office buildings, with varying certification levels (Passive House,

292   LEED Platinum, LEED Gold, and non-certified). All buildings were constructed between 2014

293   and 2020, enabling a comprehensive examination of EPG across different building typologies,

294   sustainability standards, and country conditions (developed and developing). Table 1

295   demonstrates the building projects and the information about the interviewees.

296   The interviewees, including project managers, mechanical engineers, and site managers, were

297   selected for their comprehensive knowledge of the buildings, from the design phase to being

298   operational. One participant served as the commissioning agent for two green buildings, one of

299   which was LEED Platinum-certified, with the other being expected to achieve LEED Gold

300   certification. On average, the experts had twelve years of experience in energy-efficient

301   buildings.

302 **Table 1.**

303 Information on buildings and interviewees (Source: Authors own work)

No	Building	Country	Building Type	Construction Year	Area	Interviewee No	Position	Years of experience
I	Passive House I	Germany	Residential	2019	4,009 m²	I1	CEO	34
II	Passive House II	Germany	Residential	2018	15,150 m²	I2	Project manager	21
III	Green Building I (LEED Gold)	Turkey	Headquarters	2020	45,782 m²	I3	Commissioning agent	12
						I4	Quality manager	8
						I5	Electrical technician	10
IV	Green Building II (LEED Platinum)	Turkey	Headquarters	2014	9,538 m²	I6	Project manager	8
						I7	Site manager	8
						I8	Mechanical engineer	8
						I3	Commissioning agent	12
V	Non-certified energy-efficient building I	Turkey	Educational	2017	17,030 m²	I9	Project manager	9
						I10	Mechanical engineer	8
VI	Non-certified energy-efficient building II	Turkey	Headquarters	2019	8,955 m²	I9	Project manager	9
						I10	Mechanical engineer	8

304

3.1.3 Formalization

In this step, taxonomies and the relationships between the concepts were developed using an iterative development process, as suggested by Fidan *et al.* (2011). Taxonomies represent formal hierarchical relationships between items (Pritoni *et al.*, 2021). Semi-structured interviews provided valuable information that helped us to develop the risk taxonomies and understand how different concepts interrelate. After the initial round of interviews, experts reviewed the identified risk parameters and relationships. In the second round, they evaluated the interrelations, indicated their agreement, or suggested revisions.

3.1.4 Implementation

The implementation step modeled taxonomies and their relationships using an ontology editor tool. Various ontology editors were used, including Protégé, NeOn Toolkit, SWOOP, Vitro, and Anzo for Excel in other studies. Protégé is widely used for modeling domain knowledge (Yuan *et al.*, 2018). Tah and Abanda (2011), Esnaola-Gonzalez *et al.* (2021), and Alsanad *et al.* (2019) have all used Protégé to translate their ontologies into a semantic web language. In this study, Protégé 5.5 was selected for its extensive use, free and open-source editing capabilities, stability within the ontology and Semantic Web community, and compatibility with other plug-ins (Abanda, 2011).

3.1.5 Validation

Ontology evaluation focuses on correctness and quality (Hlomani and Stacey, 2014) and is generally undertaken using verification or validation methods. The verification process ensures that the ontology is constructed correctly (Bilgin *et al.*, 2014), while validation checks whether it accurately models the real world in its application (Gruninger, 2019). Validation criteria include consistency, completeness, conciseness, expandability, and sensitiveness (Lovrenčić and Čubrilo, 2008).

It is necessary to ensure that the ontology is technically consistent and in compliance with OWL syntax for syntactic verification (Khalid *et al.*, 2023). In this study, this was tested using Pellet, an OWL-based reasoner. Later, the validation process was designed as a multi-step process so that the ontology could be tested using different sources of data at each step and enhanced until no further changes were required. A mixed-method research methodology was used to gather and analyze quantitative data, 5-point Likert scale ratings and qualitative data from interviews. Indeed, combining two methods can be more effective than using just one, providing deeper insights into research phenomena that cannot be fully comprehended through either qualitative or quantitative methods alone (Dawadi *et al.*, 2021). One aim of employing a mixed-method approach in research is to gather diverse yet complementary data on the same topic, enhancing

339   our understanding of research problems. In this way, data can be collected independently and

340   then integrated before interpreting the results (Dawadi *et al*., 2021). In our study, an article and

341   interviews were used as different data sources to validate the ontology.

342   In the first stage, an empirical article by Jain *et al.* (2020) was reviewed in detail to evaluate the

343   ontology's completeness and expandability. This particular article was selected because it

344   focused on four building types (apartment block, school, office, and hospital) and used energy

345   model calibration for performance gap assessment.

346   The second stage comprised the interviewing of six domain experts who were knowledgeable

347   about EPG in buildings. Interviews were conducted online in May 2023, each lasting one hour.

348   The proposed ontology was sent to experts beforehand for review. These experts, mechanical

349   engineers with an average of 25 years of experience (Table 2), were based in the UK (E1, E2)

350   and Turkey (E3, E4, E5, E6). All participants had at least eight years of experience in building

351   energy efficiency and were familiar with EPG issues.

352   Participants were introduced to the ontology's research aim and definition during the

353   interviews. The suggested classes and concepts of the ontology were presented in an Excel file.

354   Participants were asked to indicate the additions, removals, potential contradictions, and

355   suggestions for future development that they considered necessary. They also reviewed and

356   provided feedback on relationships between classes. At the end of the interviews, experts

357   evaluated the ontology's appropriateness, completeness, consistency, conciseness, and

358   expandability using a 5-point Likert scale. Completeness ensures that the area of interest is

359   suitably covered, while consistency checks for contradictions (Hlomani and Stacey, 2014).

360   Conciseness examines redundant or irrelevant elements (Mishra and Jain, 2020), while

361   expandability means adding new knowledge and definitions without modifying existing groups

362   (Lovrenčić and Čubrilo, 2008).

363 **Table 2.**

364 Profile of the interviewees in the validation stage (Source: Authors own work)

365

Validation Stage	Expert no	Profession	Country	Experience (number of years)
2nd Stage	E1	Mechanical Engineer	UK	13
	E2			10
	E3		Turkey	23
	E4			33
	E5			35
	E6			35

366 In the third stage, during a 1.5-hour interview, a mechanical engineer from Turkey with 46 years

367 of experience discussed the reasons for the gap and provided his feedback on the ontology. In

368 this way, different data and information sources were used to evaluate and validate the ontology.

369 This will be explained in detail in section 4.

370 *3.2 Tool development stage*

371 The ontology can be utilized by other researchers to develop tools tailored to specific needs. An

372 illustrative example of such a tool is provided in the article. The tool was developed using

373 Microsoft Excel Version 2406 (2024) and Microsoft Visual Basic for Applications (VBA), an

374 internal programming language used across various Microsoft applications. VBA allows users

375 to create forms with command buttons, option buttons, text boxes, scroll bars, and more,

376 enabling data entry and automated task execution. Using the tool, project stakeholders can not

377 only enter details related to their building stock, including geographical conditions, but also

378 evaluate the magnitude of the risks, and store and share this information with other project

379 stakeholders.

380 **4. Research findings**

381 This section presents the research findings from the ontology development stage, covering the

382 conceptual model, taxonomy, developed ontology, and ontology evaluation. It also introduces

383 the Excel-based tool created.

384 *4.1 Conceptual model*

385 In this study, semi-structured interviews were conducted with ten building experts to validate

386 and/or revise the risks identified in the literature, explore the relationships between the risks,

387 and develop a conceptual model. For example, additional risk factors and their relationships

388 were observed using verbal data from one of the projects, an office building in Turkey, as stated

389 below:

390 "Due to flexible work arrangements during the pandemic, fewer occupants worked in
391 offices. When the building was in use, lights were off, but the heating system was still
392 operating. Occupants complained about room temperature, especially in rooms with high
393 ceilings and cafeterias. That year, the weather was unusually severe. To address comfort
394 issues, the heating system was turned on earlier, and occupants were allowed to adjust the
395 room temperature by 2°C. An occupant survey can be conducted to better understand the
396 comfort-related issues and reasons for the gap."

397 This building's heating consumption exceeded design projections, while its electricity

398 consumption was lower than anticipated. Unexpected events, such as extreme weather and the

399 Covid-19 pandemic, caused problems or limitations concerning occupant behavior and

400 activities, creating uncertainty in simulation assumptions. The expert suggested post-occupancy

401 evaluations to manage these issues.

402 Based on a synthesis of literature review findings and interviews about building projects, a

403 conceptual model comprising forty concepts and five classes was created, as shown in Figure

404 2. The model includes five groups: energy performance gap, design assumptions,

405 problems/limitations, unexpected events and changes, and project management. The design

406 assumption group includes the simulation assumptions made during the design phase, such as

407 the thermal conductivity of materials and occupancy rates. Problems and limitations, including

408 elements like design problems and budget limitations, arise during the different stages of a

409 project's life cycle, introducing weaknesses to the system. These aspects can cause unexpected

410 events and changes (i.e., changes in project stakeholders), although these may also occur

411 independently. Factors affecting the manageability of these groups are classified under project

412 management, which contains elements like stakeholder experience, communication, and

413 training. According to the model, factors in the first three categories can trigger changes in

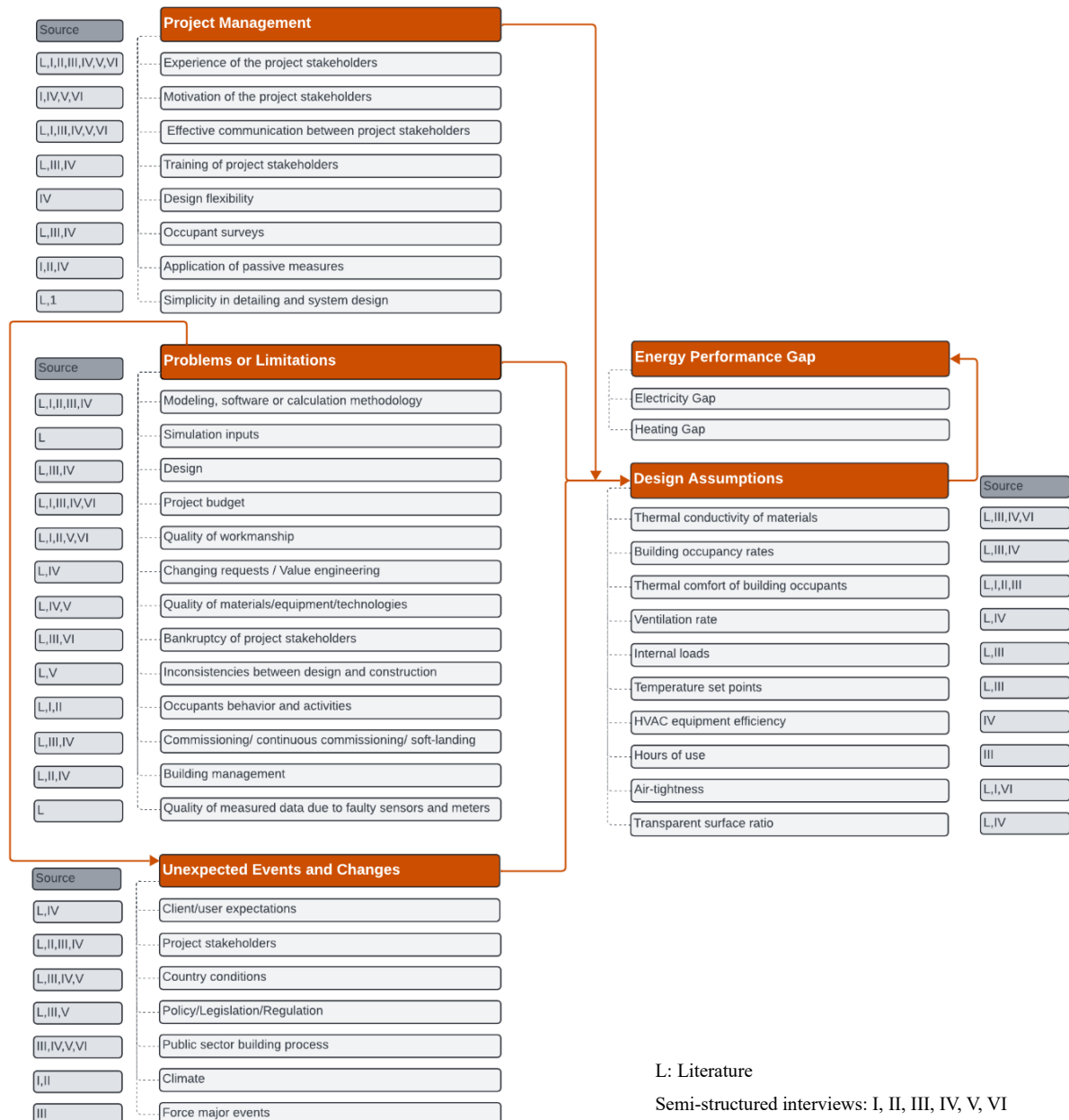414 design assumptions, leading to an energy performance gap.

Figure 2. Conceptual model (Source: Authors own work)

4.2. Taxonomy development

A taxonomy organizes elements into a superclass-subclass hierarchy. This structure brings substantial order to the model's elements, categorizes them for human interpretation, and facilitates the reuse and integration of tasks (Fidan *et al*., 2011). Figure 3 represents the taxonomy classes developed and their relationships in a Unified Modeling Language (UML) diagram. Each box represents a class and consists of three compartments in the UML diagram. The uppermost compartment contains the class name, while the middle one contains class attributes. For instance, the *Building* class has attributes such as building type, construction type, location, and project name. The relationship between the classes is shown using arrows or

426   lines. A straight line indicates an association between classes. Association role labels (e.g.,

427   "has," "results in," "causes") on the lines indicate the role of the classes. For example, the

428   Building class "has" an energy performance gap. Unexpected Events and Changes "cause"

429   Problems or Limitations, and vice versa. Multiplicities in UML diagrams indicate the number

430   of instances associated with instances of another class. For instance, multiplicity (1…*)

431   indicates that one or more Unexpected Events and Changes cause one or more Problems or

432   Limitations. While a solid line with a filled arrowhead indicates a directed relationship, a solid

433   line with an unfilled arrowhead shows inheritance between classes. For instance, the Risks class

434   is the super-class of Project Management, Problems or Limitations, and Unexpected Events and
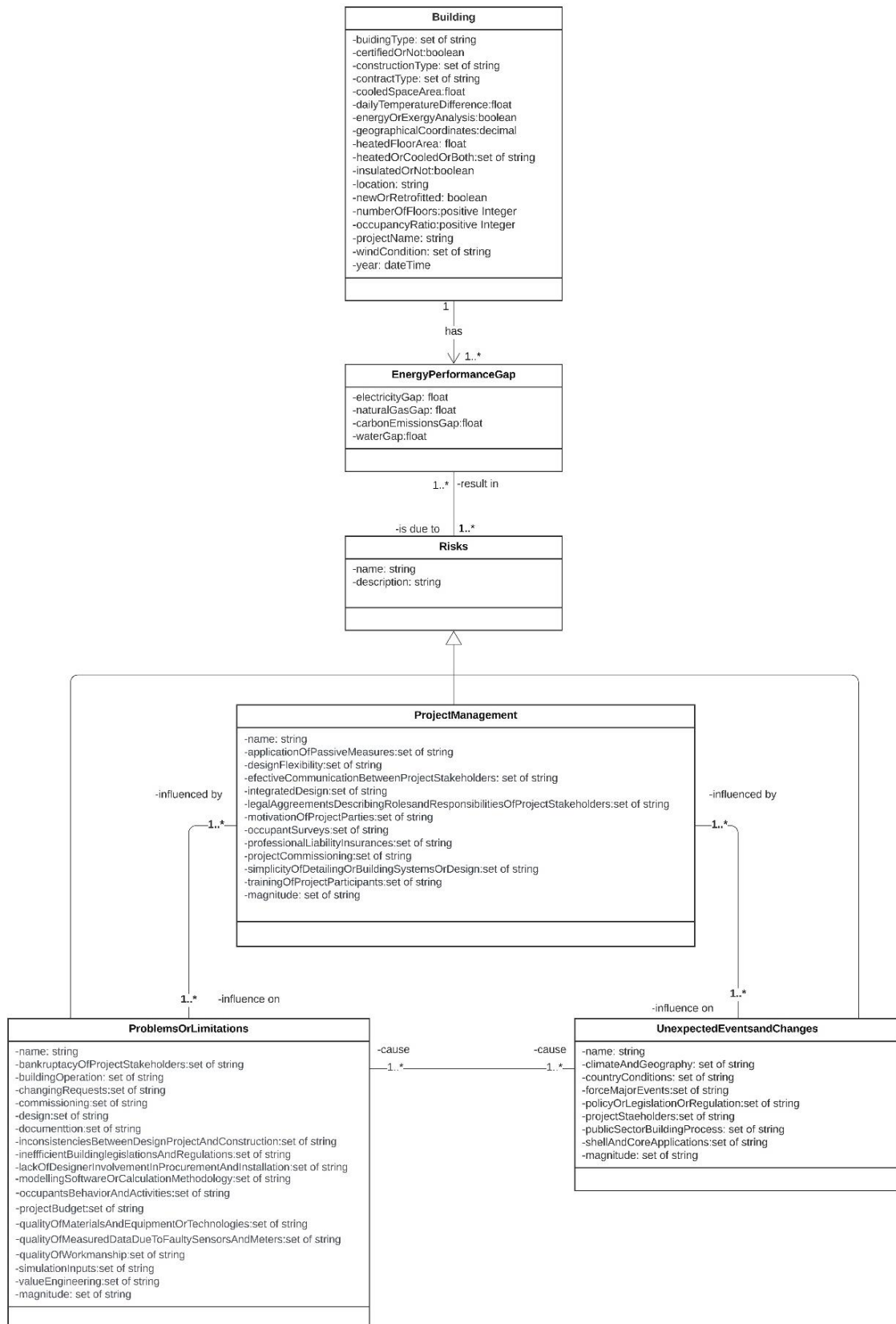
435   Changes.

Building

-buidingType: set of string
-certifiedOrNot:boolean
-constructionType: set of string
-contractType: set of string
-cooledSpaceArea:float
-dailyTemperatureDifference:float
-energyOrExergyAnalysis:boolean
-geographicalCoordinates:decimal
-heatedFloorArea: float
-heatedOrCooledOrBoth:set of string
-insulatedOrNot:boolean
-location: string
-newOrRetrofitted: boolean
-numberOfFloors:positive Integer
-occupancyRatio:positive Integer
-projectName: string
-windCondition: set of string
-year: dateTime

1

has

1..*

EnergyPerformanceGap

-electricityGap: float
-naturalGasGap: float
-carbonEmissionsGap:float
-waterGap:float

1..* -result in

-is due to   1..*

Risks

-name: string
-description: string

ProjectManagement

-name: string
-applicationOfPassiveMeasures:set of string
-designFlexibility:set of string
-efectiveCommunicationBetweenProjectStakeholders: set of string
-integratedDesign:set of string
-legalAggreementsDescribingRolesandResponsibilitiesOfProjectStakeholders:set of string
-motivationOfProjectParties:set of string
-occupantSurveys:set of string
-professionalLiabilityInsurances:set of string
-projectCommissioning:set of string
-simplicityOfDetailingOrBuildingSystemsOrDesign:set of string
-trainingOfProjectParticipants:set of string
-magnitude: set of string

-influenced by   -influenced by

1..*   1..*

1..* -influence on   1..* -influence on

ProblemsOrLimitations

-name: string
-bankruptacyOfProjectStakeholders:set of string
-buildingOperation: set of string
-changingRequests:set of string
-commissioning:set of string
-design:set of string
-documenttion:set of string
-inconsistenciesBetweenDesignProjectAndConstruction:set of string
-ineffficientBuildinglegislationsAndRegulations:set of string
-lackOfDesignerInvolvementInProcurementAndInstallation:set of string
-modellingSoftwareOrCalculationMethodology:set of string
-occupantsBehaviorAndActivities:set of string
-projectBudget:set of string
-qualityOfMaterialsAndEquipmentOrTechnologies:set of string
-qualityOfMeasuredDataDueToFaultySensorsAndMeters:set of string
-qualityOfWorkmanship:set of string
-simulationInputs:set of string
-valueEngineering:set of string
-magnitude: set of string

-cause   -cause
1..*   1..*

UnexpectedEventsandChanges

-name: string
-climateAndGeography: set of string
-countryConditions: set of string
-forceMajorEvents:set of string
-policyOrLegislationOrRegulation:set of string
-projectStaeholders:set of string
-publicSectorBuildingProcess: set of string
-shellAndCoreApplications:set of string
-magnitude: set of string

436

Figure 3. Data model for risk-energy performance gap ontology (Source: Authors own work)

438   *4.3 The developed ontology*

439   The energy performance gap-risk ontology was developed using Web Ontology Language

440   (OWL) to represent concepts, properties, and relationships. OWL is a standard language for

441   describing ontologies (Delgoshaei *et al.*, 2018). An OWL ontology includes individuals,

442   properties, and classes. Individuals, or instances, represent objects within a specific domain.

443   Classes encompass individuals, and properties are binary relations between individuals

444   (Horridge and Brandt, 2011). OWL has three types of properties: object properties, data

445   properties, and annotation properties. Object properties link individuals, data properties link an

446   individual to an XML Schema Datatype value or an RDF literal, and annotation properties add

447   more information to classes, individuals, and object/data properties (Horridge and Brandt,

448   2011).

449   The ontology consists of three main classes: Building, Energy Performance Gap, and Risks.

450   The Risks class contains three subclasses: Project Management, Problems or Limitations, and

451   Unexpected Events and Changes (see Appendix). The following sections explain the classes,

452   properties, and individuals of the ontology.

453   *4.3.1 Building class*

454   The Building class collects general information about building projects to provide a clear

455   understanding of the project's initial conditions. Concepts include Project Name, Building

456   Type, Construction Type, Number of Floors, Heated Floor Area, Certification Status, and

457   whether the building is New or Retrofitted. Object properties like "has," "has-Gap," and "has-

458   Risk-Of" link elements such as Project Name and Problems or Limitations. Data properties,

459   such as "has-Name" and "has-Number-Of-Floors," link objects to specific data types like

460   strings or positive integers. Individuals in this class include residential and non-residential

461   building types, contract types, and wind conditions.

462   *4.3.2 Energy Performance Gap class*

463   The Energy Performance Gap class includes concepts for different types of gaps, such as

464   Carbon Emissions, Electricity, Natural Gas, and Water. These gaps are linked to various risk

465   factors through object properties like "is-due-to" to define their relationships. Studies examine

466   total electricity consumption (Shi *et al.*, 2019) and gas for domestic hot water, fan electricity,

467   pump electricity, lighting electricity, and heating and cooling electricity as energy items in their

468   analyses (Chang *et al.*, 2020).

469 *4.3.3 Risks class*

470 The Risks class comprises Problems or Limitations, Unexpected Events and Changes, and

471 Project Management. Construction projects face numerous risks and uncertainties that can

472 delay completion, result in exceeded budgets, and compromise safety, quality, and operational

473 demands (Öztaş and Ökmen, 2005).

474 The Problems or Limitations subclass includes seventeen concepts (Figure 4). This category

475 lists risk factors specific to individual project phases, such as design, construction, and

476 operation, which can weaken the system and affect energy performance. For instance, poor

477 workmanship during construction can impact the building's energy performance during

478 operation. Additionally, risks throughout the project life cycle are characterized by their

479 magnitude, which can be very low, low, medium, high, or very high. The data property

480 "hasMagnitude" links an individual to a string representing this value.
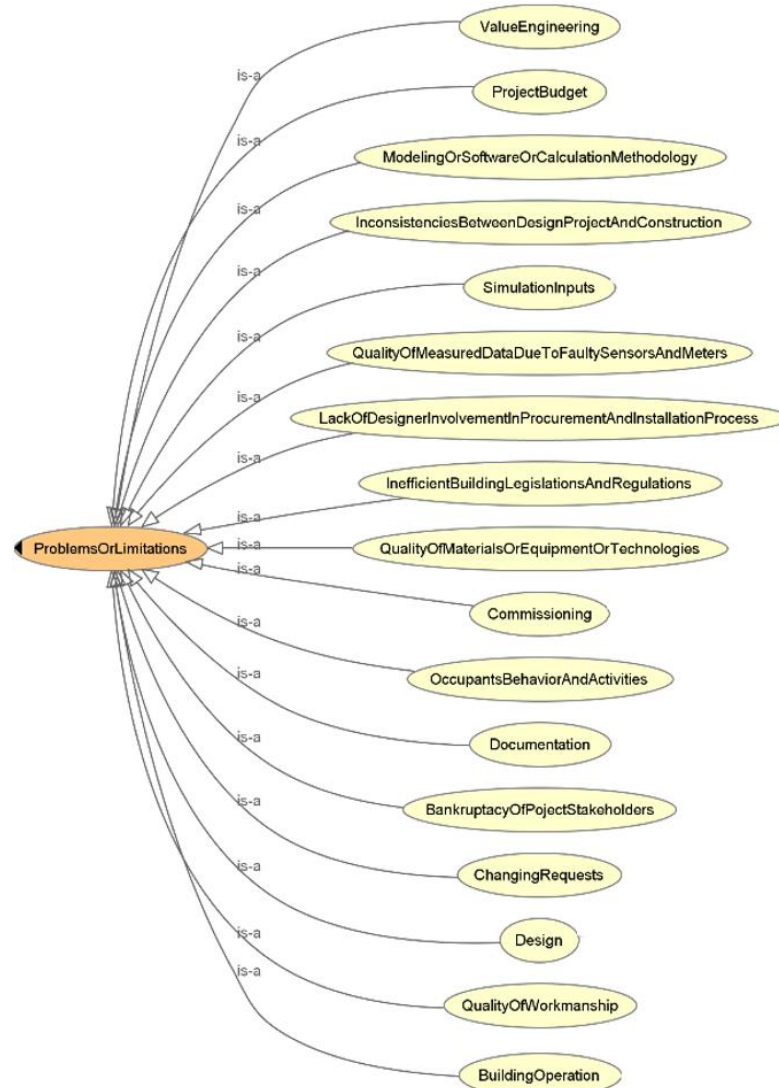
481 **Figure 4.** Problems or limitations OWLViz asserted class hierarchy

482 (Source: Authors own work)

483   Inaccurate assumptions about simulation inputs during the design phase are a primary cause of

484   the energy performance gap. The Simulation Inputs concept is categorized as a risk under the

485   Problems or Limitations class. Figure 5 lists the assumptions that can cause EPG.

486

487   **Figure 5.** Inaccurate design assumptions OWLViz asserted class hierarchy

488   (Source: Authors own work)

489   The Unexpected Events and Changes subclass contains seven concepts, while the Project
490   Management subclass contains twelve. Figure 6 illustrates the asserted class hierarchy of the
491   Unexpected Events and Changes. This subclass includes risks that cause deviations from the
492   project's initial conditions due to sudden changes and events, such as a pandemic, regulatory
493   changes, stakeholder changes, and unavailability of certified equipment. Concepts within this
494   subclass include Country Conditions, Force Majeure Events, and Climate and Geography.
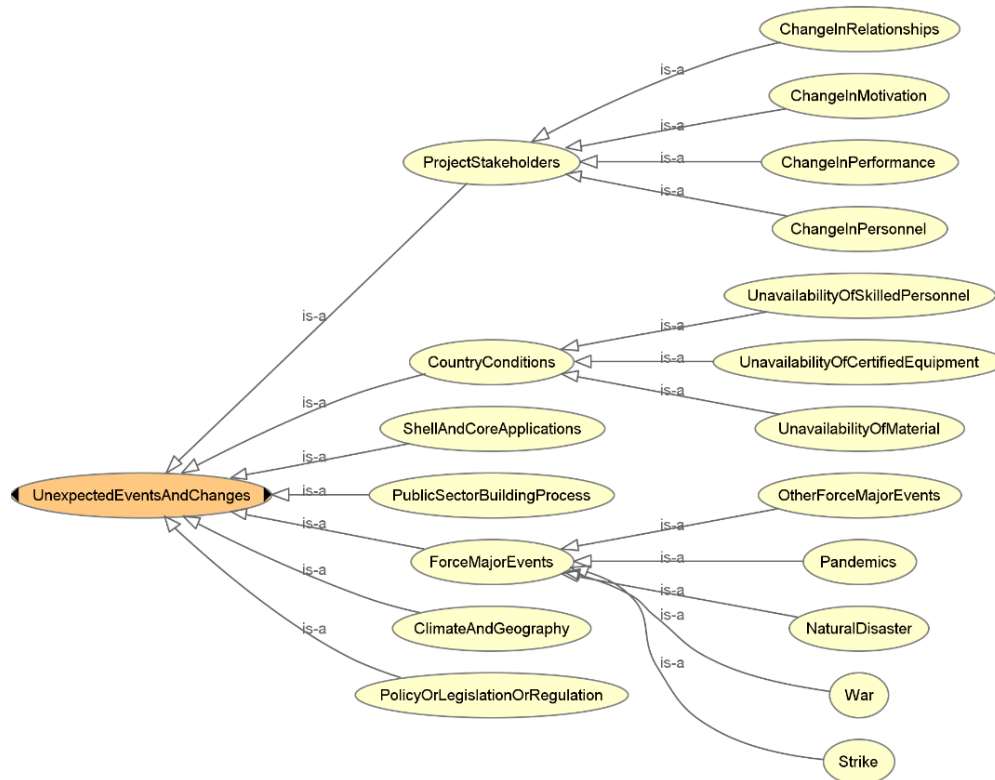
495
496   **Figure 6.** Unexpected events and changes OWLViz asserted class hierarchy (Source: Authors
497   own work)

498   The Project Management subclass includes risks that influence resilience and affect the
499   manageability of those risks causing the energy performance gap. For example, effective
500   communication between project stakeholders ensures better information flow and collaboration
501   to resolve issues across project phases. This subclass encompasses concepts such as the
502   Experience of Project Stakeholders, Integrated Design, and Design Flexibility.

503   *4.4 Ontology validation*

504   This section presents the results of the evaluation process, which included a three-stage
505   validation process.

506   In the first stage, an empirical article (Jain *et al.,* 2020) was reviewed to assess the ontology's
507   completeness and expandability. The article included four case studies, and data was manually
508   extracted to compare it with the suggested ontology. New concepts were added to the

509 appropriate class if the article mentioned a gap-causing concept not included in the ontology.

510 For example, Documentation and Poorly Specified Energy Targets were added to the Problems

511 or Limitations class and the concept of Building Management was modified to Building

512 Management and Maintenance.

513 In the second stage, interviews were conducted with six domain experts. This validation stage

514 resulted in several additions, particularly to the Buildings, Problems or Limitations, and Project

515 Management classes. For instance, Geographical Coordinates, Wind Conditions, and Energy

516 and Exergy Analysis were suggested for the Building class. Mechanical System Design

517 (including Errors in Mechanical Design Assumptions, Overdesign of Mechanical Systems, and

518 Using Incorrect Weather Data) was also recommended for the Problems or Limitations subclass.

519 Moreover, the "Design Assumptions" class, previously shown in the conceptual model (Figure

520 2), was redefined as an attribute of the "Problems or Limitations" subclass. The importance of

521 concepts such as Integrated Design, Professional Liability Insurance, and Good Interpretation

522 of Design was noted in Project Management.

523 Moreover, at the end of the interviews, six experts evaluated the ontology's appropriateness,

524 completeness, consistency, conciseness, and expandability using a 5-point Likert scale. Small

525 sample sizes are a common limitation in quantitative studies on risks in green building projects.

526 However, this constraint is understandable given the relatively smaller number of green

527 building practitioners compared to other sectors in the construction industry (Nguyen and

528 Macchion, 2023).

529 Table 3 presents the participants' responses using the mean, median, and interquartile ranges

530 (IQR). Descriptive statistics were used by Lee *et al*. (2017) and Alberici *et al*. (2020) despite

531 the sample sizes being small (six and twenty, respectively). Alberici et al. (2020) demonstrated

532 that small sample sizes can be evaluated using the median and interquartile range (IQR). The

533 median and the IQR are commonly used to assess the central tendency and dispersion of a

534 dataset. They are more robust than the mean and standard deviation because they are less

535 affected by outliers. Moreover, the IQR is particularly effective for analyzing skewed

536 distributions (Frost, 2024).

537 Experts evaluated the ontology's appropriateness, expandability, and consistency, giving it a

538 median score of 4.00 and an interquartile range (IQR) of 0.00. An IQR of 0.00 means there is

539 no variability among the middle half of the ratings. For completeness and conciseness, the

540 ontology received a median score of 4.00 and an IQR of 1.00, indicating some variability among

541 the middle half of the ratings.

542

543 **Table 3.**

544 Evaluation of the ontology based on appropriateness, completeness, consistency, conciseness,

545 and expandability (Source: Authors own work)

No.	Questions	P1	P2	P3	P4	P5	P6	Mean	Median	IQR
1	How appropriate do you think the proposed ontology is to identify the risks that cause EPG in buildings?	4	4	4	4	4	3	3.83	4.00	0.00
2	Please evaluate the completeness of the proposed ontology.	4	3	4	4	4	3	3.66	4.00	1.00
3	Please evaluate the consistency of the proposed ontology.	4	4	5	4	4	3	4.00	4.00	0.00
4	Please evaluate the conciseness of the proposed ontology.	4	3	5	4	4	3	3.83	4.00	1.00
5	Please evaluate the expandability of the proposed ontology.	2	4	4	4	4	5	3.83	4.00	0.00
*Answers to each question are given using a 5-point Likert scale.										

546 In the third stage, a mechanical engineer provided insights into the performance gap in

547 buildings. The interview highlighted several critical factors: Involvement of experienced

548 stakeholders, significance of mechanical system design, designer involvement during usage,

549 quality of commissioning, and regular equipment maintenance. This validation stage confirmed

550 that the ontology effectively captured these factors, therefore, no modifications were necessary.

551 Table 4 details the concepts added, the modifications to concept names, and their classification

552 into appropriate classes or subclasses during the validation stages.

553 **Table 4.**

554 Updates to the ontology following the validation stage (Source: Authors own work)

Stage	Type of change	Concept	New Concept Name	Sub-class	C1	C2	C3	C4	C5	C6
I	New additions	Documentation			√					
		Thermal Bridges		Inaccurate Design Assumptions						√
		Water Usage		Inaccurate Design Assumptions						√
		Poorly Specified Energy Targets		Building Design	√					
	Modification of the name	Building Management	Building Management and Maintenance		√					
II	New additions	Certified or not			√					
		Cooled Space Area			√					
		Daily Temperature Difference			√					
		Energy or Exergy Analysis			√					
		Geographical Coordinates			√					
		Heated or Cooled or Both			√					
		Number of Floors			√					
		Occupancy Ratio			√					
		Wind Condition			√					
		Year of Retrofitting			√					
		Carbon Emissions Gap							√	
		Water Gap							√	
		Hot Water Gap							√	
		Inaccurate Determination of Measurement Points		Commissioning		√				
		Incorrect Automation Algorithm		Commissioning		√				
		Building Design		Design		√				
		Mechanical System Design		Design		√				
		Errors in Mechanical Design Assumptions		Mechanical System Design		√				
		Overdesign of Mechanical Systems		Mechanical System Design		√				
		Using Incorrect Weather Data		Mechanical System Design		√				
		Lack of Designer Involvement in Procurement and Installation				√				
		Building Orientation		Simulation Inputs						
		Building Zoning		Simulation Inputs						
		Heat Losses		Simulation Inputs						
		Thermal Transmittance (Floors, Roof, and Walls)		Simulation Inputs						
		Water Usage (Cold and Hot Water)		Simulation Inputs						
		Weather Bin Data		Simulation Inputs						
		Shell and Core Applications					√			
		Integrated Design						√		
		Professional Liability Insurance						√		
		Project Commissioning						√		
		Balancing		Project Commissioning				√		
		Consideration of Occupancy Rate Afterwards		Project Commissioning				√		
		Good Interpretation of Design		Project Commissioning				√		
		Recommissioning When Necessary		Project Commissioning				√		
		Retro-commissioning		Project Commissioning				√		
		Building Maintenance		Building Operation	√					
	Modification of the name	Heating Gap	Natural Gas Gap						√	
		Building Management and Maintenance	Building Operation							
		Climate	Climate and Geography		√					
		Change in Design Assumptions	Inaccurate Design Assumptions	Simulation Inputs						√
		Changing Requests and Value Engineering	Changing Requests				√			
		Changing Requests and Value Engineering	Value Engineering				√			
III		No Changes								

C1: Building, C2: Problems or Limitations, C3: Unexpected Events and Changes, C4: Project Management, C5: Energy Performance Gap, C6: Change in Design Assumptions

555

556   *4.5 EPG-RISK tool*

557   An EPG-RISK identification tool based on Microsoft Visual Basic for Applications in Excel

558   and Macro was created using the ontology developed to demonstrate its use in practice. The

559   tool comprises seven Excel worksheets.

560   The first worksheet, ABOUT, provides users with information about the tool. The following

561   five worksheets consider the classes and sub-classes of the ontology.

562   The second worksheet, BUILDING INFORMATION, collects general data about the project.

563   Users enter energy performance gap information in the third worksheet. Data is entered

564   manually or by selecting from the dropdown menu, as demonstrated in Figure 7.

565

566

Figure 7. Building information & energy performance gap worksheet (Source: Authors own
work)

569   The fourth worksheet, PROBLEMS OR LIMITATIONS, allows users to evaluate their project

570   based on seventeen criteria, ranging from very low to very high, with an option for "not

571   applicable" (NA) responses using option boxes. This rating system allows users to compare

572   knowledge from various projects and pinpoint the most problematic criteria. Users can conduct

573   a more detailed evaluation by considering sub-criteria, such as identifying which design

574   assumptions (e.g., hours of use, airtightness, building orientation) posed more problems during

575   building energy performance calculations.

576 The fifth worksheet, UNEXPECTED EVENTS AND CHANGES, allows users to evaluate
577 their project based on seven criteria using option buttons. This section addresses various
578 unexpected conditions, such as force-majeure events like a pandemic.

579 The sixth worksheet, PROJECT MANAGEMENT, lists twelve criteria that might help to
580 control the magnitude of the gap in the project (Figure 8). Entering data for multiple projects
581 allows users to see project conditions in which a lower or higher EPG was observed.
582 Furthermore, users leverage the tool to inform their project development decisions.

583

584

585 **Figure 8.** Energy performance gap risk identification tool (Source: Authors own work)

586 Analyzing the dataset collected in the seventh worksheet (DATA) can identify where the
587 majority of projects face issues. This analysis can provide new directions for both project
588 stakeholders and policymakers to address EPG challenges in both existing and new buildings.

589 **5. Discussion**

590 *5.1 Energy performance gap-risk ontology*

591 This research standardizes experience-based and scientific knowledge on EPG in buildings by
592 developing an ontology linking risks with the energy performance gap. The ontology is crucial
593 for (1) providing linguistic unity across scientific literature and industrial practice, (2)
594 facilitating knowledge sharing among project stakeholders, and (3) enabling computer

595   readability and automatic processing in various applications. The ontology can improve

596   industry practices by facilitating risk identification, mitigation, and management.

597   The ontology developed comprises three main classes: Building, Energy Performance Gap, and

598   Risks. The Risks class is divided into three subclasses: Problems or Limitations, Project

599   Management, and Unexpected Events and Changes. Previous research on risks impeding

600   building energy performance has been limited by reliance on single case studies (Doylend,

601   2015) or literature reviews (Alam *et al*., 2017), restricting the scope to specific renovation

602   approaches (Topouzi *et al*., 2019) and the UK construction industry (Thompson *et al*., 2022).

603   Since risks vary between buildings (De Wilde, 2024), stakeholders, and countries (Yang *et al*.,

604   2016), it is essential to consider different building types, country conditions, and stakeholders

605   during risk identification. Our study addresses this gap by combining a comprehensive literature

606   review with semi-structured interviews from building projects representing various building

607   types and country-specific conditions (Turkey and Germany). Additionally, interviews with

608   architects, mechanical and civil engineers, a materials manufacturer, and an electrical

609   technician provided a multidisciplinary perspective on the ontology development. The ontology

610   identified 36 main risk factors, and 95 in total, when considering additional risks associated

611   with certain factors.

612   *5.2 Risks influencing the energy performance gap*

613   Despite using different terminologies, the literature on risk management and energy

614   performance gaps in buildings revealed many similarities with the risks identified in the current

615   ontology. Human elements, such as stakeholder communication, experience, motivation,

616   stakeholder responsibilities, occupant behavior, poor workmanship, design changes, and

617   modeling errors are prevalent in EPG. Risks also stem from poor quality materials and

618   technologies, design complexity, regulatory issues, and building maintenance. These findings

619   align with earlier research by Mahdavi & Berger (2024), Godefroy (2022), Thompson *et al*.

620   (2022), Topouzi *et al*. (2019), Gram-Hanssen and Georg (2018), Alam *et al*. (2017), Kampelis

621   *et al*. (2017), and Doylend (2015), due to the common methods used in the research.

622   The ontology development process identified new risk factors contributing to the energy

623   performance gap. For example, interviewees from two projects in Turkey, a developing

624   country, highlighted construction companies going bankrupt, which harmed construction

625   quality. Additionally, interviewees from four projects noted that the public sector building

626   process posed risks, including difficulties in selecting contractors and challenges associated

627   with using products that enhance energy performance. The lack of local, high-quality

628   mechanical equipment was also a country-specific risk in three out of four buildings in

629   Turkey. These risks affected building energy performance, construction costs, and schedule.

630   Interviewees from both Turkey and Germany expressed concerns about poor workmanship, and

631   modeling, software, and calculation methodologies. The importance of effective

632   communication and stakeholder experience was emphasized in both countries. These results

633   agree with Yang *et al*. (2016), indicating that different stakeholders and countries encounter

634   distinct risks. Consequently, it is crucial to customize risk management strategies that address

635   the specific needs and contexts.

636   The ontology helps illustrate how different factors interact to contribute to EPG. For instance,

637   project management aspects (e.g., the experience of project stakeholders) can influence

638   problems or limitations (e.g., design issues) and unexpected events and changes (e.g., those

639   related to project stakeholders) during the building life cycle. Unexpected events (e.g., a

640   pandemic) can cause problems or limitations (e.g., simulation inputs). The ontology suggests

641   that factors such as professional liability insurance, stakeholder motivation, effective

642   communication, experience, training, integrated design, simplicity of detailing, building

643   systems or design, and project commissioning can help manage EPG in buildings.

644   *5.3. Excel-based tool for energy performance gap risk identification*

645   Building on the established ontology, a tool was developed in Excel using VBA and Macros to

646   systematically collect, store, and share the risk information relating to building projects. This

647   tool may help stakeholders, such as energy service companies, project managers, energy

648   consultants, and engineers, when addressing EPG. Users can input details related to building

649   stock and geographical factors, such as construction type, number of floors, wind conditions,

650   and EPG of their projects.

651   Comprehensive project data enables researchers to uncover new insights through various

652   statistical methods. For example, Firth *et al*. (2024) identified correlations between the gap and

653   variables such as property type, floor area, year of construction, latitude, and mean gas

654   consumption. The tool also allows inputs for carbon emissions and water usage gaps,

655   broadening the scope of EPG studies beyond traditional energy performance metrics. Janser *et*

656   *al*. (2020) criticize the typical definition of EPG for often overlooking several critical aspects

657   of energy performance: greenhouse gas emissions linked to energy demand, embodied energy,

658   and the discrepancy between the optimal and planned energy performance.

659   Users can assess the magnitude of risks, which are categorized in different sheets, to help

660   prioritize certain risks and take actions to reduce the gap. Listing risks in a structured format

661   enables stakeholders to spot weak points quickly. Project teams can save information for

662   multiple projects, share it with team members, and use it as a reference for future risk

663   management. The tool essentially serves as a project risk checklist, facilitating risk

664   identification and decision support to mitigate EPG. Analyzing the collected data can pinpoint

665   common issues from different projects, offering new directions for stakeholders and

666   policymakers to tackle EPG challenges. Additionally, the collected data can be used in AI and

667   machine-learning models to develop predictive models.

668   Ultimately, the tool supports multiple stakeholders, such as industry practitioners,

669   policymakers, homeowners, and tenants in reducing the financial burden of the EPG and

670   enhancing stakeholder credibility. Moreover, by supporting more transparent and effective risk

671   management, the tool contributes to the sustainable development goals (SDG). Specifically, it

672   aligns with SDGs 11 (sustainable cities and communities), 12 (responsible consumption and

673   production), 13 (climate action), and 17 (partnerships for the goals).

674   **6. Conclusions**

675   The building life cycle involves numerous risks that complicate accurate performance

676   predictions, making effective risk identification crucial for studying EPG in buildings. Previous

677   studies have examined many factors contributing to EPG, but the disorganized handling of these

678   factors hinders efficient knowledge sharing and comparison.

679   To address these challenges, this study developed an ontology based on a literature review and

680   semi-structured interviews with industry professionals regarding six buildings in order to

681   structure concepts and factors to interrelate energy performance gap and risk in buildings. The

682   interviews helped identify new risk factors, such as stakeholder bankruptcy, public sector

683   building processes, and a lack of high-quality mechanical equipment, which are particularly

684   relevant to developing countries. Interviewees also highlighted risks related to poor

685   workmanship, modeling, software, and calculation methodologies, and emphasized the

686   importance of effective communication and stakeholder experience.

687   An Excel-based tool was created using the ontology to collect, store, and share risk data from

688   projects. This tool supports stakeholders by facilitating risk management throughout the project

689   life cycle. The tool can help reduce EPG and its financial burden on different stakeholders,

690   enhance the credibility of designers, engineers, and policymakers, and contribute to the

691   sustainable development goals through effective risk analysis. Analyzing data from multiple

692   projects can identify common issues, providing new directions for policymakers. The tool can

693   also be combined with machine learning to develop prediction models and strategies to

694   minimize EPG.

695   Although the proposed ontology was validated for its appropriateness, completeness,
696   consistency, conciseness, and expandability, the study has some limitations. These include the
697   limited number of building projects and countries involved in the ontology's development, as
698   well as the small number of experts in the validation phases. Consequently, the ontology and
699   the associated tool are mainly suitable for similar contexts, such as emerging markets in green
700   buildings, and countries with well-developed passive house construction. However, to enhance
701   generalizability, an extensive literature review has been carried out and a mixed-method
702   validation process was followed to capture the global experiences within this domain.
703   Therefore, adjustments may be necessary when using the ontology and the tool in different
704   country and sustainable building contexts. Future research using different building projects and
705   knowledge from different parts of the world may be carried out to test and improve the
706   ontology, if needed. Additionally, future research can leverage the ontology to develop new
707   tools, for example, for quantitative risk analysis, to enhance risk-based decision-making and
708   help establish more realistic energy performance targets.
709

710 **Declaration of competing interest**
711 The authors declare no conflict of interest.

References

Abanda, F.H., Kamsu-Foguem, B. and Tah, J.H.M. (2017), "BIM – New rules of measurement ontology for construction cost estimation ", *Engineering Science and Technology, an International Journal*, Vol. 20 No.2, pp.443-459.

Alam, M., Phung, V.M., Zou, P.X.W. and Sanjayan, J. (2017), "Risk identification and assessment for construction and commissioning stages of building energy retrofit projects", *Proceedings of the 22nd International Conference on Advancement of Construction Management and Real Estate*, Melbourne, Australia.

Albaret, M. and Deas, J. (2023), "Semi structured Interviews", Fanny, B. *et al.*(Ed.), *International Organizations and Research Methods: An Introduction*, University of Michigan Press, Michigan, MI, pp.82–89.

Alberici, F., Delbarba, E., Manenti, C., Econimo, L., Valerio, F., Pola, A., *et al.* (2020), "A single center observational study of the clinical characteristics and short-term outcome of 20 kidney transplant patients admitted for SARS-CoV2 pneumonia", *Kidney international*, Vol. 97 No.6, pp.1083-1088.

Alencastro, J., Fuertes, A. and De Wilde, P. (2024), "Investigating the influence of quality management on building thermal performance", *Engineering, Construction and Architectural Management*, Vol. 31 No. 8, pp. 3356-3376.

Alsanad, A.A., Chikh, A. and Mirza, A. (2019), "A domain ontology for software requirements change management in global software development environment", *IEEE Access*, Vol.7, pp.49352-49361.

Bai, Y., Yu, C. and Pan, W. (2024), "Systematic examination of energy performance gap in low-energy buildings", *Renewable and Sustainable Energy Reviews*, Vol.202, pp.114701.

Bilgin, G., Dikmen, I. and Birgonul, M.T. (2014), "Ontology evaluation: An example of delay analysis", *Procedia Engineering*, Vol. 85, pp.61-68.

Burman, E., Mumovic, D. and Kimpian, J. (2014), "Towards measurement and verification of energy performance under the framework of the European directive for energy performance of buildings", *Energy*, Vol. 77, pp. 153-163.

Calì, D., Osterhage, T., Streblow, R. and Müller, D. (2016), "Energy performance gap in refurbished German dwellings: lesson learned from a field test", *Energy and Buildings*, Vol.127, pp.1146-1158.

Chang, S., Castro-Lacouture, D. and Yamagata, Y. (2020), "Estimating building electricity performance gaps with internet of things data using bayesian multilevel additive modeling", *Journal of Construction Engineering and Management*, Vol.146 No.12, 05020017.

Corry, E., Pauwels, P., Hu, S., Keane, M. and O'Donnell, J. (2015), "A performance assessment ontology for the environmental and energy management of buildings", *Automation in Construction*, Vol. 57, pp.249-259.

Danish, M. S. S. and Senjyu, T. (2023), "Shaping the future of sustainable energy through AI-enabled circular economy policies", *Circular Economy*, Vol. 2 No.2, pp.100040.

Dawadi, S., Shrestha, S. And Giri, R. A. (2021), "Mixed-methods research: A discussion on its types, challenges, and criticisms", *Journal of Practical Studies in Education*, Vol.2 No.2, pp. 25-36.

De Wilde, P. (2014), "The gap between predicted and measured energy performance of buildings: A framework for investigation", *Automation in Construction*, Vol. 41, pp. 40-49.

Delgoshaei, P., Heidarinejad, M. and Austin, M.A. (2018), "Combined ontology-driven and machine learning approach to monitoring of building energy consumption", *Proceedings of*

Building Performance Modeling Conference and SimBuild, Chicago, United States, pp.667-674.

Doylend, N. (2015). *Evaluating building energy performance: a life-cycle risk management methodology*. Ph.D. Thesis. Loughborough University.

Esnaola-Gonzalez, I., Bermúdez, J., Fernandez, I. and Arnaiz, A. (2021), "EEPSA as a core ontology for energy efficiency and thermal comfort in buildings", *Applied Ontology,* Vol. 16 No.2, pp.193-228.

Fernández-López, M., Gomez-Perez, A. and Juristo, N. (1997), "METHONTOLOGY: from ontological art towards ontological engineering", *AAAI Conference on Artificial Intelligence.*

Fidan, D.G., Dikmen, I., Tanyer, A. and Birgonul, M. (2011). "Ontology for Relating Risk and Vulnerability to Cost Overrun in International Projects", *Journal of Computing in Civil Engineering*, Vol.25, pp.302-315.

Firth, S. K., Allinson, D. and Watson, S. (2024), "Quantifying the spatial variation of the energy performance gap for the existing housing stock in England and Wales", *Journal of Building Performance Simulation*, pp. 1-18.

Frei, B., Sagerschnig, C. and Gyalistras, D. (2017), "Performance gaps in Swiss buildings: An analysis of conflicting objectives and mitigation strategies", *Energy Procedia*, Vol. 122, pp. 421-426.

Frost, J. (2024), "Interquartile Range (IQR): How to find and use it", available at: https://statisticsbyjim.com/basics/interquartile-range/ (accessed 20 November 2024)

Garwood, T. (2019). *Closing the performance gap in building energy modelling through digital survey methods and automated reconstruction*. Ph.D. Thesis. The University of Sheffield.

Gram-Hanssen, K. and Georg, S. (2018), "Energy performance gaps: Promises, people, practices", *Building Research and Information*, Vol.46 No.1, pp.1-9.

Godefroy, J. (2022), *TM54: Evaluating operational energy use at the design stage*, Chartered Institution of Building Services Engineers, London.

Gruninger, M. (2019), Ontology validation as dialogue. CEUR Workshop Proc, 2518.

Gupta, R., Gregg, M. and Cherian, R. (2013), "Tackling the performance gap between design intent and actual outcomes of new low / zero carbon housing", *Proceedings of ECEEE 2013 Summer Study*, Toulon/Hyères, France, pp.1315-1328.

Gupta, R., Gregg, M. and Cherian, R. (2020), "Developing a new framework to bring consistency and flexibility in evaluating actual building performance", *International Journal of Building Pathology and Adaptation*, Vol. 38 No.1, pp. 228–255.

Guyo, E.D., Hartmann, T. and Snyders, S. (2023), "An ontology to represent firefighters' data requirements during building fire emergencies", *Advanced Engineering Informatics*, Vol. 56, 101992.

Han, J., Jeong, Y.K. and Lee, I. (2015), A rule-based ontology reasoning system for context-aware building energy management. *Proceedings of IEEE International Conference on Computer and Information Technology, Ubiquitous Computing and Communications, Dependable, Autonomic and Secure Computing, Pervasive Intelligence and Computing*, Liverpool, United Kingdom, pp. 2134-2142.

Hahn, J., Schumacher, P., Lang, W. and Jensch, W. (2020), Performance gap and occupant behavior - Review and analysis of high-efficiency residential buildings in Germany. *Proceedings of the 33rd International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems*, Osaka, Japan, pp.2023-2035.

Hlomani, H. and Stacey, D. (2014), "Approaches, methods, metrics, measures, and subjectivity in ontology evaluation: A survey", *Semantic Web and Information Systems*, Vol. 1 No. 5, pp. 1-11.

Hong, T., D'Oca, S., Taylor-Lange, S.C., Turner, W.J.N., Chen, Y. and Corgnati, S.P. (2015), "An ontology to represent energy-related occupant behavior in buildings. Part II: Implementation of the DNAS framework using an XML schema", *Building and Environment*, Vol. 94, pp. 196-205.

Horridge, M. and Brandt, S. (2011), "A practical guide to building OWL ontologies using Protégé 4 and CO-ODE tools edition 1.3", available at: https://mariaiulianadascalu.com/wp-content/uploads/ (accessed 10 December 2022)

Iqbal, R., Murad, M. A. A., Mustapha, A. And Sharef, N. M. (2013), "An analysis of ontology engineering methodologies: A literature review", *Research Jjournal of Applied Sciences, Engineering and Technology*, Vol. 6 No.16, pp. 2993-3000.

Jain, N., Burman, E., Stamp, S., Mumovic, D. and Davies, M. (2020), "Cross-sectoral assessment of the performance gap using calibrated building energy performance simulation", *Energy and Buildings*, Vol. 224,110271.

Janser, M., Hubbuch, M. and Windlinger, L. (2020), "Call for a definition and paradigm shift in energy performance gap research", *IOP Conference Series: Earth and Environmental Science*, Vol. 588 No. 5, pp. 052052.

Jayasudha, K. and Vidivelli, B. (2016), "Analysis of major risks in construction projects", *ARPN: Journal of Engineering and Applied Sciences*, Vol. 11 No.11, pp. 6943-6950.

Jiang, X., Wang, S., Liu, Y., Xia, B., Skitmore, M., Nepal, M. and Ghanbaripour, A. N. (2023), "A method for the ontology-based risk management of PPP construction projects", *Construction Innovation*, Vol. 23 No.5, pp. 1095-1129.

Kampelis, N., Gobakis, K., Vagias, V., Kolokotsa, D., Standardi, L., Isidori, D., *et al*. (2017), "Evaluation of the performance gap in industrial, residential & tertiary near-Zero energy buildings", *Energy and Buildings*, Vol.148, pp. 58–73.

Khalid, Q., Fernandez, A., Lujak, M. and Doniec, A. (2023), "SBEO: Smart building evacuation ontology", *Computer Science and Information Systems*, Vol. 20 No.1, pp. 51-76.

Kwaśnik, B. H. (2020), "Changing perspectives on classification as a knowledge-representation process", *Knowledge Organization*, Vol.46 No.8, pp.656-667.

Lee, W., Lin, K. Y., Seto, E. and Migliaccio, G. C. (2017), "Wearable sensors for monitoring on-duty and off-duty worker physiological status and activities in construction", *Automation in Construction*, Vol. 83, pp. 341-353.Liang, J., Qiu, Y. and Hu, M. (2019), "Mind the energy performance gap: Evidence from green commercial buildings", *Resources, Conservation and Recycling*, Vol. 141, pp. 364-377.

Lovrenčić, S. and Čubrilo, M. (2008), "Ontology evaluation- Comprising verification and validation", *Proceedings of 19th Central European Conference on Information and Intelligent Systems*, Varazdin, Croatia.

Madhusanka, H.W.N., Pan, W. and Kumaraswamy, M.M. (2022), "Stakeholder engagement and collaboration in overcoming the constraints to delivering low carbon buildings in high-rise high-density cities". *In IOP Conference Series: Earth and Environmental Science*, Vol. 1101 No. 4 pp. 042030.

Mahdavi, A. and Berger, C. (2024), "Ten questions regarding buildings, occupants and the energy performance gap", *Journal of Building Performance Simulation*, pp. 1–11.

852 Menezes, A. C., Cripps, A., Bouchlaghem, D. and Buswell, R. (2012), "Predicted vs. actual
853   energy performance of non-domestic buildings: Using post-occupancy evaluation data to
854   reduce the performance gap", *Applied energy*, Vol.97, pp. 355-364.

855 Mills, E., Kromer, S., Weiss, G. and Mathew, P.A. (2006), "From volatility to value: Analyzing
856   and managing financial and performance risk in energy savings projects", *Energy Policy*,
857   Vol. 34, pp. 188-199.

858 Mishra, S. and Jain, S. (2020), "Ontologies as a semantic model in IoT", *International Journal*
859   *of Computer Applications*, Vol. 42 No.3, pp. 233-243.

860 Moradi, S., Hirvonen, J. and Sormunen, P. (2024), "A qualitative and life cycle-based study of
861   the energy performance gap in building construction: Perspectives of Finnish project
862   professionals and property maintenance experts", *Building Research & Information*, Vol. 52
863   No.5, pp.564-576.

864 Nguyen, H. D. and Macchion, L. (2023), "A comprehensive risk assessment model based on a
865   fuzzy synthetic evaluation approach for green building projects: the case of
866   Vietnam", *Engineering, Construction and Architectural Management*, Vol.30 No.7,
867   pp.2837-2861.

868 Ortiz, M., Itard, L. and Bluyssen, P.M. (2020), "Indoor environmental quality related risk
869   factors with energy-efficient retrofitting of housing: A literature review", *Energy and*
870   *Buildings*, Vol. 221, 110102. Öztaş, A. and Ökmen, Ö. (2005), "Judgmental risk analysis
871   process development in construction projects", *Building and Environment*, Vol. 40 No. 9,
872   pp.1244-1154.

873 Pomponi, F. and Moncaster, A. (2018), "Scrutinising embodied carbon in buildings: The next
874   performance gap made manifest", *Renewable and Sustainable Energy Reviews*, Vol. 81, pp.
875   2431-2442.

876 Pritoni, M., Paine, D., Fierro, G., Mosiman, C., Poplawski, M., Saha, A., *et al.* (2021),
877   "Metadata schemas and ontologies for building energy applications: A critical review and
878   use case analysis", *Energies*, Vol. 14 No. 7, pp. 1-37.

879 Qin, X., Mo, Y. and Jing, L. (2016), "Risk perceptions of the life cycle of green buildings in
880   China", *Journal of Cleaner Production*, Vol. 126, pp. 148-158.

881 Saunders, M., Lewis, P. and Thornhill, A. (2019), Research methods for business students,
882   Pearson Education, London, England.

883 Schachinger, D. and Kastner, W. (2017), "Ontology-based generation of optimization problems
884   for building energy management", *Proceedings of IEEE International Conference on*
885   *Emerging Technologies and Factory Automation*, Limassol, Cyprus, pp. 1-8.

886 Shelbourn, A.M., Bouchlaghem, D.M., Anumba, C.J., Carillo, P.M. Khalfan M.M.K. and
887   Glass, J. (2006), "Managing knowledge in the context of sustainable construction", *Journal*
888   *of Information Technology in Construction*, Vol. 11, pp. 57-71.

889 Shi, X., Si, B., Zhao, J., Tian, Z., Wang, C., Jin, X., *et al.* (2019), "Magnitude, causes, and
890   solutions of the performance gap of buildings: A review", *Sustainability*, Vol. 11, pp. 937.

891 Shrubsole, C., Hamilton, I.G., Zimmermann, N., Papachristos, G., Broyd, T., Burman, E., *et al.*
892   (2019), "Bridging the gap: The need for a systems thinking approach in understanding and
893   addressing energy and environmental performance in buildings", *Indoor and Built*
894   *Environment*, Vol. 28 No.1, pp. 100-117.

895 Siraj, N.B. and Fayek, A.R. (2019), "Risk identification and common risks in construction:
896   Literature review and content analysis", *Journal of Construction Engineering and*
897   *Management*, Vol. 145 No. 9, 03119004.

Tah, J.H.M. and Abanda, H.F. (2011), "Sustainable building technology knowledge representation: Using semantic web techniques", *Advanced Engineering Informatics*, Vol. *25* No.3, pp.547-558.

Thompson, D., Burman, E., Mumovic, D. and Davies, M. (2022), "Managing the risk of the energy performance gap in non-domestic buildings", *Building Services Engineering Research and Technology*, Vol.43 No.1, pp.57-88.

Topouzi, M., Killip, G., Fawcett, T. and Owen, A. (2019), "Deep retrofit approaches: managing risks to minimise the energy performance gap", *Proceedings of the ECEEE 2019 Summer Study on Energy Efficiency*, Presqu'ile de Giens, France, pp.1345-1354.

Tserng, H.P., Yin, S.Y.L., Dzeng, R.J., Wou, B., Tsai, M.D. and Chen, W.Y (2009), "A study of ontology-based risk management framework of construction projects through project life cycle", *Automation in Construction*, Vol. 18 No.7, pp. 994-1008.

Van Dronkelaar, C., Dowson, M., Spataru, C. and Mumovic, D. (2016), "A review of the regulatory energy performance gap and its underlying causes in non-domestic buildings", *Frontiers of Mechanical Engineering*, Vol. 1 No.17.

Varpio, L., Paradis, E., Uijtdehaage, S. and Young, M. (2020), "The distinctions between theory, theoretical framework, and conceptual framework", *Academic Medicine*, Vol. 95 No.7, pp. 989-94.

Wang, X., Yuan, J., You, K., Ma, X. and Li, Z. (2023), "Using real building energy use data to explain the energy performance gap of energy-efficient residential buildings: A case study from the hot summer and cold winter zone in China", *Sustainability*, Vol.15 No.2. pp. 1575.

Yang, R.J., Zou, P.X.W. and Wang, J. (2016), "Modelling stakeholder-associated risk networks in green building projects", *International Journal of Project Management*, Vol. 34 No.1, pp. 66-81.

Yuan, J., Li, X., Chen, K. and Skibniewski, M.J. (2018), "Modelling residual value risk through ontology to address vulnerability of PPP project system", *Advanced Engineering Informatics*, Vol. 38, pp.776-93.

Yousri, E., Sayed, A.E.B., Farag, M.A. and Abdelalim, A.M. (2023), "Risk identification of building construction projects in Egypt", *Buildings*, Vol.13 No.4, pp. 1084.

Zheng, Z., Zhou, J., Jiaqin, Z., Yang, Y., Xu, F. And Liu, H. (2024), "Review of the building energy performance gap from simulation and building lifecycle perspectives: Magnitude, causes and solutions", *Developments in the Built Environment*, pp. 100345.Zhao, X., Hwang, B.G. and Gao, Y. (2016), "A fuzzy synthetic evaluation approach for risk assessment: A case of Singapore's green projects", *Journal of Cleaner Production*, Vol. 115, pp. 203-213. Zhou, P. and El-Gohary, N. (2017), "Ontology-based automated information extraction from building energy conservation codes", *Automation in Construction*, Vol. 74, pp. 103-117.

Zhou, Z., Goh, Y.M. and Shen, L. (2016), "Overview and analysis of ontology studies supporting development of the construction industry", *Journal of Computing in Civil Engineering*, Vol. 30 No.6, pp.1-14.

Zou, P. X. W., Xu, X., Sanjayan, J. and Wang, J. (2018), "Review of 10 years research on building energy performance gap: Lifecycle and stakeholder perspectives", *Energy and Buildings*, Vol.178, pp.165-181.
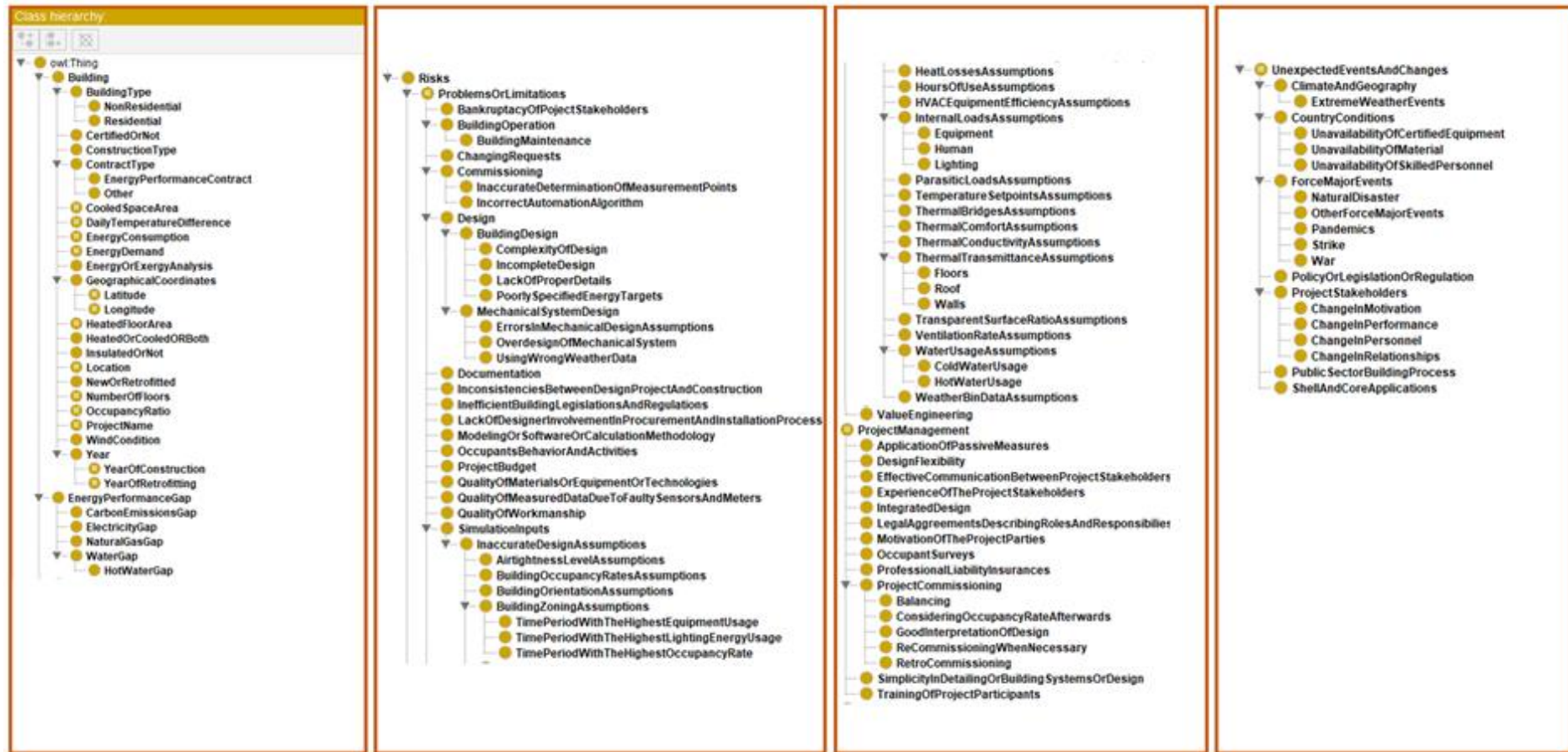
940 **Appendix**

941
942 **Figure A1.** Classes of the energy performance gap-risk ontology (Source: Authors own work)