The Role of Semantics in Enhancing User Experience in Building and City Web Applications

Thesis for Doctor of Philosophy (Engineering) (Jul) (RFPDEGND)

Andrei Hodorog

June 2023

Contents

Abstract

This thesis embarks on an exploratory journey through Building Information Modelling (BIM) and City Information Modelling (CIM) within web applications, aiming to significantly uplift citizen engagement and satisfaction. At its core, the thesis proposes an innovative framework that uses semantics to intricately weave contextual information into user experience (UX), fostering innovative applications tailored to the built environment. The intellectual pursuit is meticulously structured around three pivotal research questions, each unfolding a distinct but interconnected inquiry stage.

The first research question delves into Enhancing Learning UX with Semantics. It seeks to uncover how semantics can amplify the learning experience within existing web applications. This stage is marked by the development of a semantic web-based mining environment meticulously designed to unravel and map the intricate web of roles and skills pivotal in BIM. The endeavour goes beyond mere identification; it strategically establishes correlations, paving the way for learning pathways tailored and resonant with the evolving dynamics of the built environment.

Progressing to the second stage, the thesis casts its investigative net into Context Derivation in Smart Cities. This stage is not just about exploring methods but pioneering ways to extract context from the rich tapestry of static and dynamic artefacts embedded within a Digital Twin framework. The goal? To elevate the UX in smart city applications to unprecedented heights. This stage is characterised by the strategic leveraging of BIM semantics, with the aim of transforming the user experience of a diverse cohort of stakeholders, ranging from architects and urban planners to engineers. It is an endeavour that transcends the conventional, blending advanced methodologies to enrich interactions within the web of smart city ecosystems.

The journey culminates with the third research question, which focusses on Semantic Scaling and Social Media Analysis. This stage is visionary in its approach, envisioning the scaling of semantics at the city level and positioning citizens as active sensors in an ever-evolving urban landscape. The ambition is grand – to develop a taxonomy model rooted in a semantic-based risk model. However, the thesis does not stop there; it ventures into the vibrant world of social media data streams. By applying sophisticated natural language processing (NLP) techniques, research meticulously sifts through digital chatter, aiming to uncover hidden narratives that weave together environmental factors, risk events, and the pulse of citizen satisfaction.

The findings of this thesis are not only insightful; they are transformative. The research demonstrates the practical applicability of semantics across three core dimensions. In socio-organisational aspects, the thesis sheds light on the dynamic nature of construction skills, underscoring the imperative for adaptive training methodologies that keep pace with the rapid evolution of BIM roles. The exploration does not stop at the micro level; it extends its gaze to the macro-grain of the built environment. The thesis showcases the profound impact of advanced web technologies, such as the VueJS front-end framework and innovative web builders. When these technological marvels are harmoniously integrated with core UX principles, they unravel complex phenomena, weaving a tapestry of enhanced UX within the pulsating heart of smart cities.

The thesis also pioneers social media analytics, presenting it as a formidable information source that can significantly shape smart city decision making. The insights gleaned are not

just data points; they are statistically significant revelations that empower stakeholders, offering them the clarity and foresight to make decisions that are not just informed but visionary.

As such, this thesis is not just a scholarly endeavour, but a beacon that illuminates the path for future explorations and developments. It is a testament to the synergistic fusion of information science techniques and smart city communities, significantly contributing to the rapidly evolving landscape of semantic integration and UX enhancement within the built environment. The journey embarked on in this thesis is not just about answering questions; it is about charting new territories, opening new horizons, and setting the stage for a future where the built environment is not just smart, but sentient, responsive, and perpetually in tune with the needs and aspirations of its citizens.

ACL Access Control List

AI Artificial Intelligence

API Application Programming Interface

AR Augmented Reality

BAS Building Automation System

BIM Building Information Modelling

BMS Building Management Systems

BRE Building Research Establishment Ltd

CCTV Closed-Circuit Television Cameras

CDM Data Model Schema

CPU Central Processing Unit

CRES Metropolia University of Applied Sciences and Center For Renewable Energy Sources

CSTB Centre Scientifique et Technique du Bâtiment

CSS Cascading Style Sheets

CUSP Computational Urban Sustainability Platform

DCMS Department for Digital, Culture, Media & Sport

DDL Data Definition Language

DT Digital Twin

FEA Finite Element Analysis

GDPR General Data Protection Regulations

GIS Geographical Information Systems

GUI Graphical User Interface

HCI Human-Computer Interaction

HPC High-Performance Computing

HTML HyperText Markup Language

ICT and Communication Technologies

IFC Industry Foundation Classes

IoT Internet of Things

ISO International Standards Organisation

IT Information Technology

JSON JavaScript Object Notation

KIS Keep It Small

LDA Latent Dirichlet Allocation

LIST Luxembourg Institute of Science and Technology

LSTM Long Short-Term Memory

MIB Make It Big

MQTT Message Queue Telemetry and Transmission

MR Mixed Reality

MRA Multiple Regression Analysis

MVP minimum viable product

NIST National Institute of Standards and Technology

NLP Natural Language Processing

OSINT Open-Source Intelligence

OS Operating System

OWL Ontology Web Language

PDF Portable Document Format

RDBMS Relational Database Management System

REST Representational State Transfer

SE Standard Error

SPARQL SPARQL Protocol and RDF Query Languag

SSN Semantic Sensor Networks

SVD Singular Value Decomposition

TF-IDF Term Frequency - Inverse Document Frequency

UCD User Centric Desing

UK United Kingdom

UN United Nations

URI Uniform Resource Identifier

UX User Experience

XML Extensible Markup Language

1 Introduction

The Web has transformed the way we access and manage information, leading to a wealth of data across various sectors. In architecture, engineering and construction (AEC), efficient data management is essential to enhance design, sustainability, and project coordination [211]. However, the volume and complexity of this data present significant challenges [6].

The proliferation of interconnected devices has resulted in an exponential growth of data, presenting new opportunities and challenges for the management of information [46]. In this context, the emerging field of the Semantic Web aims at structuring and annotating data to define meaning and enable more intelligent computing [158], transforming the current state of the Web from a vast collection of loosely connected documents into a well-integrated knowledge base, where data is equipped with machine-readable metadata. This shift promises to improve search capabilities and enable logical inferences, which are particularly beneficial in the management of complex datasets prevalent in the AEC industry [280].

Concurrently, as the Semantic Web transforms the foundations of data management, complementary technological developments are enabling new approaches in specific domains. A new framework known as Building Information Modelling (BIM) has emerged that provides machine-readable representations of buildings and their properties throughout the project lifecycle [368]. Despite its advances, integrating BIM with the dynamic aspects of urban development remains a challenge [290]. As such, the integration of semantic web techniques presents immense potential to leverage BIM's digital information meaningfully. This thesis investigates the intersection of Semantic Web technologies and BIM and explores how they can collaboratively enhance Web applications, energy management, and smart city infrastructure, providing a more structured and intelligent approach to information management in the AEC domain [280].

The impact of the Internet extends beyond technology; it catalyses global change, addressing critical issues such as climate change and urban sustainability. As noted by Castells [70], Hilty et al. [155], and Grimm et al. [134], the Internet has significantly influenced the built environment, intersecting societal demands and environmental concerns. The Semantic Web's potential to improve user experience (UX) in web applications by clarifying complex data from the built environment is yet to be fully leveraged, as discussed by Bibri and Krogstie [45] and Wang et al. [374].

This thesis posits that semantics can significantly improve UX by enabling a meaningful interpretation of this data. Furthermore, the research contends that BIM should evolve to encapsulate physical structures and dynamic human interactions within these spaces. The underlying motivation lies within the fact that by expanding the scope of BIM to include dynamic semantic models, we can gain invaluable insight into user behaviour and sentiment on a city-wide scale [77].

Based on these gaps and opportunities in urban informatics, this research presents a specific hypothesis, objectives, and questions to address. The thesis places itself at the forefront of this challenge, providing a critical analysis of the role of semantics in elevating UX and an examination of persistent challenges within the built environment. In addition, it presents a detailed exploration of semantic solutions and data analytics, adopting an interdisciplinary approach to efficiently manage urban assets.

1.1 Research aim and objectives

This thesis aims to investigate the potential of semantic web technologies to improve decision-making processes and user experience (UX) in web applications pertaining to building and urban data management.

More specifically, this thesis focuses on leveraging semantic web techniques and building information modelling to develop ontology-driven web interfaces, optimise energy management in buildings, and create interconnected smart city networks.

The objectives of the research are the following:

1. To analyse current learning mechanisms in the construction industry and develop a semantically-enhanced platform to improve training and upskilling related to BIM adoption.

2. To design and implement a semantics-based user interface for an energy management platform that derives context from static and dynamic artefacts to optimise usability.

3. To develop an urban-scale digital framework that utilises citizen data as a semantic layer for extracting insights about events, environmental factors, and satisfaction.

4. To assess the role of semantics in elevating user experience across individual buildings, energy systems, and city-wide networks through user studies and expert feedback.

The research also explores getting context from digital twins that represent both static buildings and dynamic artefacts. and also investigates applying these ideas at a larger scale, from individual buildings to entire cities.

The overarching hypothesis that informs this research is that **BIM Web applications can be enhanced through the use of a Digital Twin, using context derived from static artefacts (buildings) and dynamic artefacts in a way that maximises User Experience.**

To evaluate this hypothesis and contribute to the body of knowledge, the following key research questions will be addressed:

1. How could an existing web application be enhanced using context to maximise user experience?

2. How to derive context from static artefacts (buildings) and dynamic artefacts using a digital twin?

3. How can we rely on this diverse (i.e., tangible and intangible) dynamic context to enhance a user's experience (e.g., a stadium, city planner, city manager) in various applications in the built environment?

The first research question examines how semantic context can be integrated into existing web applications to enrich the user experience. The second question explores the derivation of the context through digital twins that represent both static buildings and dynamic systems. The third question investigates applying this approach on multiple scales, from individual buildings to city-wide networks.

As such, the research questions delve into semantic integration approaches, leveraging digital twins for context derivation, and scaling these techniques from buildings to cities. Answering these questions will generate valuable knowledge regarding the role of semantics in enhancing user experience in the built environment.

By studying these questions through analysis, building prototypes, and testing with users, this research seeks to show how semantics and digital twins can improve web applications for buildings and cities. It focuses on making web interfaces, energy systems, and smart cities better by adding structure and meaning to their information.

The goal is to demonstrate interfaces that are easier to use, buildings that are more energy efficient, and cities that use citizen knowledge. This can lead to new theories and techniques to apply semantics for improved user experience and smarter cities.

1.2 Contributions to the body of knowledge

The contributions of this research are interconnected and build upon each other in various ways.

Firstly, the research developed models for integrating semantic context into web applications to enhance user experience. The context was instrumental in creating a semantic search feature specifically tailored for the construction industry. By leveraging the theoretical techniques of contextual integration, this feature significantly enriched learning and training resources within the industry.

Second, the thesis extended its focus to using digital twins and semantic modelling. This combination was pivotal in deriving contextual information from both static and dynamic artefacts. The insights gained here informed the participatory design of an energy management platform. This platform benefited from user feedback to refine its user interface and experience and was fundamentally underpinned by the synergies between the digital twin and semantic modelling theories.

The third area of contribution involved scaling semantic frameworks. Research methodologies evolved from focusing on individual buildings to encompassing city-wide networks. This scaling was critical in transforming citizens into active sensors within these networks. It also facilitated the natural language processing of social media data, thus integrating citizen knowledge into smart city networks using expanded semantic frameworks.

Finally, the research converged the semantic web, Building Information Modelling (BIM), and ontological modelling methodologies. This convergence was crucial in elevating decision-making processes across various applications and scales. It guided a multidisciplinary approach that linked all areas of the research. This holistic approach laid the theoretical foundation for a smart city prototype. Combining semantics, digital twins, and AI, this prototype represents a significant advancement in urban planning and management, exemplifying the practical application of the multidisciplinary theoretical approach adopted by the thesis.

1.3 Limitations of the research

While this research makes several notable contributions, it is also essential to acknowledge its limitations. Firstly, the research primarily focused on incorporating semantics to enhance user experience within specific contexts like BIM web applications, energy management platforms, and smart city data. Additional work is needed to assess the applicability of the approaches across other systems and domains.

Second, the evaluation of the improved user experiences was based on relatively small sample sizes of users and experts. More extensive empirical studies are needed to validate the effectiveness of semantic enhancements for broader populations of users. Relatedly, the research was concentrated in particular geographic areas and cultural contexts. Testing user experience improvements in diverse demographics and locales could strengthen the conclusions.

In addition, rapidly evolving technologies pose a challenge in terms of the longevity and relevance of proposed solutions. While the research methodology was designed to adapt to technological changes, long-term studies are needed to ascertain the robustness of the frameworks over time. From a practical standpoint, adopting the approaches relies on stakeholder participation and overcoming inertia around existing legacy systems, which requires further analysis.

Finally, while precautions were taken to protect user data and privacy throughout the research process, increasing concerns around data ethics warrant an ongoing review of the data collection, retention, and sharing protocols followed in the research. Expanding the oversight and audit mechanisms could improve alignment with emerging privacy standards. This research provides substantial evidence for the integration of semantics to improve user experience within BIM, energy management, and smart city systems. However, the highlighted limitations indicate several worthwhile directions for further research to build on these contributions. Addressing these limitations could strengthen the validity, applicability, and responsiveness of the approaches explored, providing fertile ground for future research.

1.4 A summary of Building Information Modelling, Semantics and User Experience

To provide conceptual grounding, this section summarises key literature on BIM, semantics, and UX.

By converging BIM, semantics, and digital twin technologies, smart city initiatives could be revolutionised to transcend conventional city planning and management approaches, fostering urban areas that are not only practical and innovative, but also environmentally sustainable. This anticipatory transition will be critical to improving the quality of life of urban residents, thus cementing the value and urgency of the research presented in this thesis [396, 48].

This structured research aims to provide empirical evidence on how semantics-driven UX in BIM can redesign the software landscape for the betterment of urban living. Research anticipates revealing how next generation web-based BIM tools can become more effective by highlighting the criticality of UX in the broader context of smart city development and the sustainable evolution of the built environment [154].

1.4.1 Enhancing User Experience Through the Integration of BIM and Semantics

Building on the smart city concepts previously discussed, using Building Information Modelling (BIM) with extra detailed information is vital in improving User Experience (UX) in architecture and construction. BIM brings a digital change, making it easier to handle much information throughout the life of a building project. This new approach is crucial for those involved in the project, as it makes designing and building more efficient by combining different types of data [48].

Adding detailed descriptions to BIM is essential to help people understand the complex parts of a building better. This extra detail clarifies the details, leading to better decisions and improving the user experience [291]. The combination of BIM and these descriptions affects everything from the structure's details to the overall project, moving towards designs that focus more on the user. This combination also leads to better teamwork and fewer errors due to better information sharing [5].

As such, combining BIM and technologies that provide detailed semantic descriptions is a major step forward in UX, providing professionals with a complete set of tools to create and understand complex construction projects.

1.4.2 The influence of semantics and Big Data on User Experience Design

As semantic enrichment aims to enhance built environment information systems, a critical complementary priority is ensuring these systems provide optimal user experiences. This highlights important considerations for UX design and evaluation in the context of BIM and smart cities.

Delving deeper into the aforementioned synergy, the semantic dimension, central to linguistics and concerned with meaning interpretation, stands as a pivotal influence in UX design. Effective semantic design ensures unambiguous comprehension of product functionality, which is essential for a frictionless user-product interaction [89]. When intricately woven into the user experience, semantic elements ease user navigation by elucidating product usage without the need for lengthy discovery processes [148]. Moreover, semantic web technologies sophisticate this landscape by facilitating the assemblage of customised data, strengthening data-driven decision making, a core principle in the context of smart cities and intelligent infrastructure [276, 182].

In the big data arena, semantic techniques increase user engagement and intelligence within civil engineering and BIM. The inclusion of synonymous and related terms within the search parameters exemplifies a semantic-based BIM strategy that enriches UX and broadens the informational horizon for users [104]. Moreover, interpreting user intent semantically curates more germane search results, thus further refining UX than their non-semantic counterparts do [22].

Semantics are crucial to delineating user expectations and facilitating navigation within digital platforms [24]. Furthermore, the concept of BIM being bolstered by semantics is not isolated. The emergence of Urban Data Lakes (UDL) and semantic technologies expands the connective tissue between BIM models and broader systems [122, 120, 165]. This multitude of connections is manifest in the accessibility of documents produced by BIM technologies that have eliminated the dependence on specialist software for online readability [103, 128].

Manufacturers and suppliers shall cooperate to establish BIM-compatible data standards that promote superior data interoperability and accessibility. Semantic technologies excel in this domain, interlacing a rich tapestry of information that is beneficial in collaborative efforts [214]. Furthermore, the relationship between semantics and BIM is evident in the context of the data supply chain, where semantics facilitate seamless data exchange through standardised interfaces [181, 94, 280].

Semantic BIM implementations, grounded in ontological principles, thus enable the creation of adaptable and networked data systems, underscoring their increasing pertinence within collaborative construction platforms [215, 180].

1.4.3 The challenges of User Experience in the design of BIM and CIM web applications

Understanding the niche of smart cities, with its unique technological and design complexities, reveals a multidimensional challenge for UX designers. Creating user interfaces that elegantly merge with advanced city infrastructures requires an understanding that extends beyond software: a deep understanding of urban planning and social behaviour [190, 152, 389, 16, 166, 384].

Designers must strive to maintain a balance between digital conveniences and the intuitive use of physical spaces, ensuring that smart city environments cater seamlessly to their inhabitants. Here, consistency in design, be it through colour coding or a structured information hierarchy, becomes paramount in reducing cognitive load and improving interaction with BIM applications [167, 314, 275].

Understanding users, their motivations, and their interactions with the application is critical. Designers should leverage interviews, surveys, and market analyses to align the product with user objectives, while navigating the complexity of designing universally comprehensible UX across various platforms [337, 325].

The UX challenges presented by smart city initiatives are numerous, ranging from physical space design to intricate interweaving of diverse data streams. The key lies in ensuring ubiquity in application accessibility, aligning with global UX guidelines, and adopting an inclusive design philosophy that takes into account non-native users [9, 361, 245, 257].

Emerging interaction paradigms, such as voice commands, offer a glimpse of the future of user participation in smart city applications, marking the path for more accessible and efficient user interfaces [373, 233]. At the same time, the responsibility to maintain clarity and order within the interfaces of these BIM applications cannot be overstated [256, 118].

Moving to the nuanced domain of Big Data, UX designers grapple with presenting colossal datasets, fostering usability, and circumventing information overload. Big data requires a user interface that not only accommodates multifaceted datasets but also makes them palpable for various demographic data of users [131].

Heavy datasets, such as those integral to BIM applications, burden computational resources and user cognition alike, slowing information retrieval and potentially impeding decision-making processes [188]. The antidote to these predicaments may lie in the realm of advanced data visualisation techniques and machine learning, which can equip users with insightful representations of the underlying datasets [367].

Understanding user interactions with data becomes a linchpin. UX designers must dive into user behaviour analytics to craft designs that reflect emerging usage patterns [227]. Consequently, UX narratives must be informed by comprehensive user research and competitive analysis, ensuring that design considerations remain both anticipatory and responsive to user requirements and market dynamics [325].

Although designing interfaces for complex urban systems poses many UX challenges, they underscore the growing need for human-centred research on semantics in the built environment. This leads to defining the scope of this thesis research.

1.5 Motivations and background of the researcher

The impetus behind this research stems from the background of the researcher that spans information security, user experience design, and web development for environmental projects. His interest in this topic started with a bachelor's degree in Information Security and Forensics. After graduating, the researcher began to specialise in user experience (UX) and information security. Before starting the Ph.D., the researcher gained industry experience as a front-end web developer at various companies.

The researcher's work gradually moved into ontologies through interface and UI development, leading him to participate in research projects like WISDOM, BIMEET, and CUSP.

The researcher worked on the WISDOM project, developing a web interface that showed water network pipes using different colours. This interface also displayed pipe coordinates fetched from an ontology via SPARQL and highlighted the water flow while marking essential parts like pipes, catchments, and subcatchments.

The researcher is interested in "smart cities" and Web interfaces for citizen engagement with local governments. The researcher believes that cities can be made more livable and sustainable by making them "smarter" through computing capabilities and aiming to create interfaces that improve communication between residents and city governments.

The researcher's academic interests align with developing more intelligent and environmentally sustainable cities using technologies like the Semantic Web. This foundation inspired the author to pursue doctoral studies on improving human-centric design and decision making for buildings and cities through semantic techniques. The research aims to create more livable, sustainable and responsive urban environments by taking a human-centred approach to semantics, user experience, and smart systems.

The researcher's mixed industry and academic professional trajectory fostered a multidisciplinary perspective spanning security, usability, ontology modelling, urban planning, and computing. This background enables a holistic approach to improving processes, interfaces, and intelligence in the built environment by elevating the role of citizens and the semantic context.

The diverse expertise drove the enthusiasm of the researcher for interweaving semantics and emerging technologies to transform physical and digital experiences in the built environment. This research aims to empower stakeholders through meaningful interactions, insights, and participatory design.

1.6 Thesis structure overview

This thesis begins with a conceptual foundation, laying a scholarly framework that extends through an exploration of smart city research, the creation of ontologies, and software development, concluding with a demonstration of the NLP work product.

Chapter 2 commences with a comprehensive review of the literature examining the symbiosis of UX and semantics, followed by an exposition of the seminal theories that underpin the role of semantics in BIM applications.

Chapter 3 delineates the research design under the lens of epistemological and philosophical foundations, detailing the iterative methodology that transcends from semantic enhancement in education to city-scale applications.

Subsequent chapters 4, 5, and 6 categorise the investigative findings based on the stage-specific methods applied. The final chapters bind the narrative, contrasting the research findings with extant literature, and crystallising the thesis's contribution to current practices en route to the conclusion summarising significant insights and outlining avenues for future research. The chapter concludes with introspections on the research domain, delimiting scholarly achievements while envisioning future directions for continuing research inquiries.

This thesis aims to bridge the gaps within the existing corpus of knowledge, challenging conventions, and pioneering novel paradigms in UX, BIM, smart city development, and semantic-driven urban evolution.

2 Literature review

This chapter provides a comprehensive review of the academic literature relevant to the research questions outlined in the Introduction. Specifically, it examines existing work on the use of semantic web technologies and modelling approaches to enhance user experience in built environment applications.

User experience (UX) refers to the perceptions, emotions, attitudes, and responses that a person has to say about using or anticipating the use of a product, system, or service. It involves an overall assessment of the usability, functionality, and design of the system based on how well it meets the person's needs, expectations, and goals within their specific context of use [207].

Semantics refers to the study of meaning, which encompasses the interpretation of words, phrases, and symbols. As explained by Cruse [89], semantics examines the relationship between linguistic units and the concepts they represent. The foundational theories of semantics posit that words and expressions have inherent meanings based on conventional usage [222]. However, context also plays a pivotal role in shaping meaning, as emphasised by scholars such as Firth [121] in his theory of syntagmatic relations.

The review commences by delineating the theoretical foundations of the semantic web and its connection to enriching digital interfaces and systems. Core concepts, technical architecture, and research trajectories are explored to situate this paradigm within the broader domains of knowledge representation and human-computer interaction.

Building on these fundamentals, the following subsection analyses the seminal and contemporary literature on user experience design, evaluation, and modelling. Theoretical theories of cognitive load, information architecture, and interaction design are reviewed to contextualise considerations to enhance user experience through semantic approaches. Specific considerations for building information systems and smart city contexts are highlighted.

Transitioning away from user experience theories, the following subsection provides an in-depth overview of building information modelling (BIM), a pivotal knowledge framework for the built environment. In this section, technical underpinnings, capabilities, limitations, and research trends are investigated to elucidate the challenges and opportunities for convergence with semantics.

Building Information Modelling (BIM) is a process that involves the generation and management of digital representations of the physical and functional characteristics of places. BIM uses three-dimensional real-time dynamic building modelling software to increase productivity in building design and construction. The model contains digital information on geometry, spatial relationships, geographic information, as well as quantities and properties of the components of the building [344].

Next, the emerging digital twin concept is explored as a bridge between BIM and broader smart city applications. The existing literature on virtual-physical environments, dynamic data modelling, and digital urban twins is synthesised to reveal research gaps at this nexus of technologies and applied contexts.

A Digital Twin is a virtual representation of a physical object or system that replicates its key characteristics and behaviours, integrating multidisciplinary data including 3D models, sensor data, machine learning algorithms, and simulations to create a living digital profile of its physical counterpart. Digital twins enable various functions such as predictive analysis, operational optimisation, simulation of real-world scenarios, effective decision making, and collaboration between stakeholders [348].

A smart city uses information and communication technologies (ICT) and the Internet of Things (IoT) to collect multi-domain data and apply advanced analytics to optimise the efficiency of urban operations and services and improve the quality of life for citizens. Integrates ICT into city infrastructure such as transportation, energy, healthcare, water, and waste management. Through citizen engagement and real-time data analysis, smart cities aim to improve sustainability, resilience, innovation, and governance [45].

The final sections examine the intersection of semantics, BIM, digital twins, and user experience within broader smart city research and implementations. Case studies, technical approaches, and design challenges are reviewed with a focus on improving urban decision making and experiences, laying the foundation of a conceptual arc from core theories to advanced applications.

This chapter provides a structured examination of the literature that illuminates the pathways to more intuitive, intelligent and human-centric built environments through the synthesis of semantic web techniques, user experience principles, information modelling, and urban informatics. The insights gained situate this thesis and reveal critical areas for contribution.

2.1 Operational concepts Introduction

As digital interfaces and experiences continue to evolve, a critical complementary priority is to ensure that these systems provide optimal user experiences. This highlights important considerations for UX design and evaluation in the context of BIM and smart cities.

To establish a conceptual foundation before surveying the landscape of existing literature, the next section will define some key terms used extensively throughout this research, including digital twins, semantic web technologies, ontologies, and linked data.

2.1.1 Semantics general technologies

The emergence of the semantic web, as envisioned by Berners-Lee et al. [41], was based on the core semantic theories to develop a framework for sharing machine-readable data with embedded semantics on the World Wide Web. Ontologies represent a key element of the semantic web, providing formal definitions of concepts and their relations within a domain, allowing a structured representation of knowledge [140, 340]. The most prominent ontology languages include RDF Schema [60] and OWL [238], which facilitate semantic reasoning and inference.

To implement semantics on the Web, markup languages such as XML [59] and microformats [187] encode metadata to provide information on the content of the page. Search engines can then take advantage of this semantic markup to improve search accuracy [142]. Web services employ common semantics to exchange data, enhanced by standards like OWL-S [231], allowing interoperability. In general, semantic web principles and technologies aim to transform the Web from a collection of linked documents to a meaningful and machine-understandable knowledge base.

2.1.2 Semantics and usability

Semantics improve understandability by providing additional contextual information and meaning to the content. As explained by Sabou [313], semantic annotations and metadata allow systems to process information in a more meaningful way that aligns with human understanding. For example, ontologies define formal conceptual structures that clarify the relationships between different terms and entities [270]. Furthermore, semantic reasoning can infer additional knowledge from explicitly stated facts, allowing a more comprehensive understanding of the information of the system [26].

According to Razmerita [298], semantics improve usability by enabling personalisation, allowing systems to tailor interfaces and services to users' specific needs and contexts. Castaneda et al. [69] elaborates that semantic representations allow mapping user profiles and preferences to appropriate content and functionality. Furthermore, Ploennigs et al. [289] explains that semantics improve interoperability between various systems and data sources. This seamless data integration provides users with a unified view and interaction, avoiding the need to adapt to different interfaces and terminology.

Semantics also facilitate findability, which enhances usability according to Morville [251]. By clarifying terminology and providing linked contextual data, semantic search can provide more accurate and complete results [142]. As Davies et al. [98] discusses, semantically annotated content can improve navigation and information retrieval by matching the intent of the user with greater precision. In general, semantics improve understandability and usability through contextual knowledge, personalisation, system interoperability, and findability. However, usability can suffer if semantic systems are too complex or poorly designed, as cautioned by Shadbolt et al. [328]. Therefore, balancing expressivity, complexity, and cognitive load is essential.

2.1.3 Ontologies

Ontologies are formal conceptual models that represent the knowledge of a specific domain through definitions of key concepts and relationships. They provide an abstract representation of entities and their properties, attributes, and interrelationships to enable the sharing and reuse of knowledge across applications. Ontologies include machine-interpretable definitions and axioms that enable automated reasoning and inference of new knowledge [340].

2.1.4 HTML5 and schema language

This subchapter explores various aspects, including the role of semantic markup, such as HTML5 and schema.org, in structuring and enhancing the meaning of web content [51]. In addition, the researcher will investigate how the application of semantic technologies, such as linked data and ontologies, can contribute to sentiment analysis in product design and create a more personalised and context-aware web experience within social networks.

Semantic markups, such as HTML5 and schema.org, play a crucial role in structuring and enhancing the meaning of web content within the Web User Experience (UX). These markup languages provide a standardised way to represent and annotate the semantics or meaning of information on the Web, enabling improved understanding, interpretation, and accessibility for humans and machines.

HTML5, the latest version of the Hypertext Markup Language, introduced new semantic elements and attributes that allow developers to create well-structured web documents [40] using semantic elements, such as “¡header¿”, “¡nav¿”, “¡section¿”, and “¡article¿”. Designers can provide more precise indications of structural components and relationships within the content. This structural clarity enhances the user’s comprehension and navigation of the web page, leading to a more meaningful and coherent user experience.

In addition to HTML5, schema.org provides a collaborative effort of the major search engines, including Google, Bing, Yahoo!, and Yandex, to establish a shared vocabulary to mark structured data on the Web [143]. Schema.org provides several semantic types, properties, and enumerations, allowing web developers to specify the meaning and context of their content more clearly. Using Schema.org annotations enriches web content with metadata, enabling search engines to understand and display information more effectively. This structured representation not only aids in search engine optimisation, but also enhances the relevance and usability of web content for users. Therefore, semantic markup languages facilitate more effective communication among content creators, users, and automated systems. By improving the semantics of web content, HTML5 and Schema.org contribute to a more accessible, interpretable, and semantically enriched web environment, thus improving user experience.

2.1.5 Sentiment analysis in product design

Within user experience (UX), the connection between information schemas, such as *schema.org*, and sentiment analysis lies within their complementary roles in understanding and enhancing the user’s perception and interaction with digital content. Although information schemas provide a structured representation of web content, enabling more apparent meaning and context, sentiment analysis employs natural language processing (NLP) techniques to analyse user-generated content and extract sentiments and emotions.

An example of an experience in which both stimulus and response are required is the interaction between the end user and the GUI (Graphical User Interface) software or any other computer-generated content. Sentiment analysis can be used to analyse the thoughts and emotions of an end-user during this interaction. This technique is particularly effective at lifting the affective state that the user is experiencing.

Lou and Yao [218] is amongst the first pioneers to describe natural language processing (NLP) as being used to characterise entities such as products, services, organisations, individuals, issues, events and other attributes, as well as attributes from other sources derived from raw material.

Although there are unresolved problems in natural language processing (NLP), one of its success stories is sentiment analysis, as it only requires that a computer understands *positive* or *negative* sentiments and the topic of each sentence. Cambria et al. [65] defines *sentiment analysis* as requiring some core components, including, but not limited to, polarity classification, mood identification, emotion recognition, agreement detection, subjectivity analysis, and identifying the degree of positivity or negativity.

Schouten and Frasincar [324] describes the levels at which sentiments can be analysed: the "document" level, the "sentence" level, and the "aspect" level. Aspect-level sentiment analysis focuses on finding sentiment-target pairs (i.e. entities or key characteristics) about the target (aspects) in a text. Tsytsarau and Palpanas [353] defines four primary methods for finding sentiment-target pairs: machine learning, dictionary-based determinations, statistical analysis, and semantic analysis.

Several methods have been used to detect general sentiment from end users, including keyword detection, lexical affinity, and statistical modelling. Keyword spotting involves classifying texts based on the presence or frequency of specific affective words, allowing the identification of sentiment-related content [281]. Lexical affinity approaches focus on measuring the similarity or relatedness of words to specific sentiment categories, providing a more nuanced analysis [130]. In this scenario, statistical modelling uses machine learning techniques to analyse substantial datasets and predict sentiment based on learnt patterns and features [397].

Despite their respective strengths and weaknesses, these approaches have been widely used in sentiment analysis. However, the keyword-spotting approach stands out due to its simplicity, accuracy, and cost-effectiveness. Focusing on specific affective words or expressions, keyword spotting provides an expeditious and straightforward assessment of sentiment without the need for complex language processing or extensive training data. This simplicity makes it particularly suitable for applications requiring real-time analysis or quick insights.

Within the context of the widely used thought-aloud protocol in product design experiments, people use language to express their thoughts about their experiences. User views about using the same product vary on the basis of several factors, including previous experience. Several paradigms support the idea that user experiences and emotions are stored in long-term memory. Two notable examples are the "affect heuristic theory" of Slovic et al. [334] and the "somatic marker theory" of Marg [229]. Emotion is a reminder of previously stored information during retrieval. Product interactions with users will evoke emotional responses based on the "affective tag" previously associated with similar situations.

Munezero et al. [252] distinguishes between "sentiments", "opinions", "feelings", "affect", and "emotions" in terms of academic terminology. The authors argue that, contrary to "emotions", which are brief and spontaneous, "sentiments" represent emotional provisions formed toward an object or virtual entity characterised by greater intensity and longer persistence.

Although the experience of meaning (usability) and the experience of emotion (affect) are critical factors when measuring user experience, empirical research has been minimal when trying to find measures for both factors. This is particularly important because the effect could enormously impact the user's perception of usability. This is described by Setchi and Asikhia [327], who proposed a method that uses image schemas, sentiment analysis, and computational semantics to evaluate the user experience. Semantic and affect analysis is used to determine words related to a particular task and words associated with a particular image schema to identify (i) task-specific words and

(ii) emotion-related words. Most importantly, the most vital contribution of their work was that it established links between image schemas and various interactions and states of mind. Their uniqueness comes from their study, which required a domain-specific ontology of image schemas specially created for this research. The ontology-based algorithm was used to identify and link the emotions and sentiments expressed by the user with the image schemas used during a specific task.

Johnson [176] defines an *image schema* as an interaction dynamic that describes the experience of interacting with the physical world, whose arousal occurs due to interactions before language and context. To determine what "up-down orientation" means, the author asks us to visualise objects that have a vertical orientation in the perception, such as trees, standing upright, climbing stairs, visualising a flagpole, measuring children's height, and experiencing water rising in a bathtub. These various experiences are examples of the abstract structure of the verticality schema. Over time, a relationship develops between the recurring experiences and what the user sees. The recurring experience of being able to compare oneself to others is "extended metaphorically" by these associations.

Fernandez et al. [119] argue that interactions in which the user persists for prolonged periods can develop interesting patterns of interaction with minimal cognitive effort executed in a shorter time.

Hurtienne et al. [168] gives examples of such habits, including the experience of visualising water rising in a river, the perception of the spatial situation at the centre or periphery of a scene, and visualising the location of an object, inside or outside a container. When end users describe these experiences, they act as a doorway to the brain's neural network, which stores the corresponding information. Examples of coded word patterns, such as "up/down" and "centre/periphery" demonstrate how image schemas were developed from repetition and similar patterns of interaction with the world, and these images leave cerebral footprints.

Hampe and Grady [146] believes that each container image schema consists of three parts: an inside, an outside, and a boundary that regulates movement within and outside the container. In other words, the container image schema is described by words such as "in, "out", "enter", "emerge", and "come out". On the other hand, Johnson [176] based the up-down image schema on gravity experiences coded in words such as "up", "down", "top", "increase", "decrease", "rise", and "fall". Verbal and non-verbal stimuli have been extensively used in cognitive linguistics to investigate several image schemas, such as "up/down", "big/small", "near/far", and "left/right". These studies demonstrate that sensory modalities activate image schemas in sight, sound, and movement. It is not universally accepted that image schemas should be classified using relationships between cognitive domains. However, researchers agree that they can be organised according to their relationships with cognitive domains.

2.1.6 Semantic Analysis in Social Media

Having explored the various methods used in sentiment analysis to detect general sentiments, such as keyword spotting, lexical affinity, and statistical modelling, the focus now shifts to another vital aspect of sentiment analysis: semantic analysis in social networks. Using the power of semantic analysis, researchers and practitioners can gain deeper insight into nuanced meanings and sentiments embedded within user-generated content on social networks, enabling a more comprehensive understanding of public opinion, trends, and user experiences.

In 2019, the use of social networks was well established and more than 70% Americans are registered with one or more of these services [72]. Multiple technical factors explain this trend. One of them is the increasing Internet coverage around the world. Almost 99% of European households had access to a 4G connection in 2018, while it was five times lower in 2012 [101]. Another factor is the widespread use of smartphones, since around 80% of the worldwide population has access to this technology [100]. These technical developments encourage more and more people to use social networks, allowing them to react to events in real-time.

Social networks represent an increasing amount of information that is of great interest to Big Data specialists. For example, on Twitter, each message contains valuable meta-data, such as the location from which the tweet was sent. It is, of course, inconceivable to manually process this amount of information. However, the very concise and standardised way of posting on Twitter, at most 280 characters, makes it easier to deploy automated models to aggregate and process this information.

There is abundant literature exploring the application of text mining techniques in various domains, including policy making [19, 259, 18]. Chai et al. [74] demonstrated that the most significant advantage posed by the integration of text mining processes into the policy making process is to alleviate the problem of bounded rationality, a long-established policy problem [239, 18], described as a core principle for incremental policy change [109]. Text mining applications are based on the adaptation principle, creating time-efficient data collection and information processing processes. This approach increases the time that decision makers have to adapt to the challenging environment and produces better results estimates. A significant example are the early warning signals that real-time text mining applications can issue, such as the prediction of earthquakes from social media content [316].

Gao and Eldin [126] developed a methodology to extract text to locate employment credentials information relevant to construction workforce professions, such as knowledge areas, skills, and expertise, from online sources of publicly available employment information. The authors used search engines (market innovations at the time) to understand the job market and employers' expectations. A system algorithm based on statistically valuable pattern extraction was performed on the qualifications selected previously and subsequently used to detect the presence of such qualifications on new pages. Once the qualification was determined, the LDA (Latent Dirichlet Allocation) model could identify the skill groups employers require. After aggregating the ten most relevant keywords from each of the ten most relevant topics, the authors concluded that adding more industry skills to the existing curriculum can help the next generation of students better secure their career paths in the construction industry.

Forman et al. [124] attempted to minimise human effort in quantifying problems in customer support centre call logs arising from unstructured free text fields. This research attempted to establish an accurate quantification aggregating expenditures broken down by the type of problem addressed to optimise human resource allocation by appropriately targeting the right engineers and providing the most effective means of identifying, assisting, and recording the most frequent problems. Manual classification requirements are eliminated through underlying approaches such as new text clustering methods, machine learning through interactive classifier training, and category quantification.

[332] has done longitudinal research over a decade to determine progress in human resource management in hospitality and identify emerging trends. Data were collected and analysed using a text mining algorithm and human judgements. The results of the content analysed by the computer and generated through human judgement were integrated and conceptually graphed into a map readable by both computers and humans. This research resulted in nine main topics in human resource management, which were later interpreted based on the relevant timestamp and geographic region.

Other social media mining algorithms explored by numerous researchers, such as Lopez-Castroman et al. [217] and Song et al. [336], focus mainly on suicide prevention, crime forecasting, and alert systems.

Other researchers have attempted to detect events in smart city contexts from social media messages. In 2013, Khan et al. [186] proposed a graph-based retrieval algorithm. They performed a latent Dirichlet allocation (LDA) to find topical clusters, pick representative tweets from clusters, and construct a lexical graph to identify key discussions. In 2014, Truong et al. [352] used a feature extraction-based approach to find relevant tweets over conversational tweets in the context of natural disasters. Combining their features with a bag-of-words approach improved the accuracy of the classification task. Nalluru et al. [255] published a multimodal model that can be tuned to detect relevant tweets using text and images. Relevance can be defined with respect to the information needs of emergency management agencies. Both text and images are converted to vectors using neural networks and singular value decomposition (SVD). Vectors were concatenated and fed into a LightGBM classifier model [255].

Although designing interfaces for complex urban systems poses many UX challenges, they underscore the growing need for human-centred research on semantics in the built environment. This leads to the necessity of exploring the meaning and fundamental operational concepts of User Experience, which are discussed in the next subsection.

As digital interfaces and experiences continue to evolve, a critical complementary priority is to ensure that these systems provide optimal user experiences. This highlights important considerations for UX design and evaluation in the context of BIM and smart cities.

The following sub-chapter delves into the impact of semantics on the design and evaluation of web interfaces. By examining the relationship between semantics and user experience, this study aims to uncover insights into how the meaning and representation of information on the Web influence user perception, cognition, and engagement.

With the key concepts defined previously, the following section will explore the historical origins and chronological development of digital twins and associated technologies to provide background context.

2.2 Digital Twins Historical Evolution

This subsection delves into the transformative potential of digital twins in redefining the traditional Human-Computer Interaction (HCI) paradigm. Digital Twins enable dynamic and interactive representation of physical objects, systems, and processes using advanced technologies such as the Internet of Things (IoT), Artificial Intelligence (AI), and real-time data integration.

2.2.1 Smart city ontologies

Digital Twins reshape the UX landscape by providing immersive, context-rich, and responsive user experiences that bridge the gap between physical and digital realms, opening new possibilities for enhanced decision-making, design optimisation, and interactive simulations.

Regardless of the approach undertaken, the high degree of complexity and the voluminous characteristics of the datasets involved in creating Digital Twins call for creating a context for the various data types and their associated real-world entities. The context of Digital Twins can be provided through semantics, which encompasses ontologies as building blocks at their technical forefront. Due to their physics-based ontology, these analogues are named and the "type" of an item is inferred.

Regarding extensive city-scale modelling, the CityGML standard [135] replaces Industry Foundation Classes (IFC) as the go-to structure due to its popularity as a standard information model. One of the first high-profile instances of its use was to semantically describe a Berlin neighbourhood for the Energy Atlas Berlin project [199]. Additional contextual information, such as building function and weather data, can produce forecasts of heating demand and solar PV availability at the building level on a city scale using the semantic decomposition of city walls, floors, and roof surfaces [75].

An urban scale ontology was also developed as part of the SEMANCO project, which included similar building-level information but also aimed to describe the infrastructure and systems for delivering electricity, mapping the end use of electricity to the energy sources and carriers [86]. The ontology was then used to create an integrated platform that allowed urban planners to perform a city-scale energy analysis and perform targeted interventions.

The SEMANCO project developed a significant smart city ontology in OWL DL-LiteA, resulting in 592 classes to achieve data integration. However, due to the lack of sensor and dynamic data provision, the SEMANCO ontology is intended to exchange data during urban planning phases. Therefore, it could contribute to a high-level ontology that links ontologies to each vertical that addresses these issues because each industry owns highly specific semantic data.

The building-level concepts of IFC and the city and district-level concepts of CityGML must be linked to manage energy holistically. An attempt was made to capture district-level energy concepts in an EE-district ontology. For interoperability at different levels, the ontology is mapped and integrated with existing upper-level vocabularies [157].

IBM created the SCRIBE smart city ontology [196], citing a lack of available ontologies and tools. Their ontology laid the foundation for establishing formal descriptions of city services, events, metadata, and abstractions.

Nesi et al. [258] created the "Knowledge Model 4 City" (also known as "KM4city") smart city ontology, which emphasises public transportation and mobility.

Consoli [85], created another smart city ontology, which combines Geographical Information Systems (GIS) data into a lightweight ontology, and proposed a linked open data modelling approach to managing smart city data.

Smart cities monitor various parameters such as energy, water, and transportation management. In addition, these cities try to prevent criminality and organise the appropriate response to an event [254]. Information and communication technologies (ICT) play a vital role in a city's ability to respond as quickly as possible to the needs of citizens [33].

As the digital economy grows and technology advances, digitalisation is critical to enabling the IoT in smart cities. A digital city is connected to the Internet. As described by Su et al. [341], it retains the necessary platforms, such as remote sensing, global positioning systems, and geographic information systems, all of which contribute to the development of the urban geographic information infrastructure required for public service. Physical products are still required. They have become "hybrid", as evidenced by smart washing machines, doorbells, and automobiles. Smart cities emerge due to a host of intelligent solutions implemented in all sectors of society, driven by a mixture of disruptive technologies and social innovations [115].

Several publications have identified three critical layers of a smart city: *perception*, *network*, and *application* [341, 212]. Internet of Things (IoT) devices collect data at the *perception* layer via sensors, RFID, and QR codes. The *network* layer provides a conduit for the perception layer's data to be exchanged and transferred via information and network management centres. The *application* layer utilises fuzzy recognition, cloud computing, and other intelligent technologies to analyse and process data. The application layer enables rapid real-time responses to physical data.

Due to the extensive use of interconnected sensors and microcontrollers, the security of smart cities is continuously debated. Kitchin and Dodge [192] investigated the impact of smart city initiatives and highlighted a paradox. These cities effectively counteract uncertainty and urban risks, but create new vulnerabilities and threats. Therefore, researchers usually include a reflection on the security of the proposed solution.

2.2.2 User Experience in a smart city context

When investigating the implications of social networks on smart cities, it is vital to understand the steps some pioneering cities have taken to build smart cities [110]. A significant amount of evidence points to the fact that the UK government and government officials use social media platforms for accountability in governance, which also helps officials to understand several events more productively [387]. Researchers in Sweden have identified that the use of social networks has proven to be much more effective in relating to the specifics of the local community. Therefore, social networks are also believed to have promoted accountability and transparency within society [38].

Mobile technology has played a critical role in making governments more efficient. Thus, in this agreement, the lead actors are used to understand the city's situation and to facilitate connections and interactions. In this agreement, the Angel Network is treated as the most similar component [223]. According to the research findings, different applications can help the public understand their mindset, which is a valuable factor for government officials to use when making decisions [309].

Regarding the ideology and conceptualisation of smart cities, the current generation of users is more drawn to the positive aspects of social networks. As a result, it can be deduced that a well-functioning government strategy promotes cooperation and collaboration between citizens and outside organisations [299]. This approach decreases the number of government-sponsored projects and, at the same time, increases the level of responsibility that citizens have in that area. Social media usage in smart city concepts has increased significantly among stakeholders. Thus, this increment is due to the administration's openness and transparency. Therefore, better administration has been aided by the use of social networks [329].

A smart city's institutional framework can be enhanced by encouraging citizen participation and improving e-participation. Those who live in smart cities are much more likely to be involved in the decision making process [271].

Many researchers have been attracted to the implementation of social networks for smart city governance, and this study examines the results of these elements under this agreement. Innovative governance and stability of electronic media must be achieved in all smart cities. In this agreement, political management must be a highly sensitive entity [38]. Many things could be improved with a right turn in this agreement. As problems in today's society are of particular importance [223], it is vital to ensure that people are adequately informed in smart cities and that the necessary level of mutual trust between citizens and government institutions is established.

Crowd-sensing applications provide context awareness capabilities for various roles in the system (e.g., citizens as users, developers, and stakeholders) and facilitate interactions between them. Hu et al. [164] shows the feasibility of crowd sensing using cellular networks to predict, transmit information, and potentially act on various types of queue that restrict the mobility of citizens in a congested city area. Although the collected data could be easily fed into a web-based Digital Twin platform for intelligence and active citizen engagement, this approach would be highly intrusive to user privacy due to real-time location monitoring requirements.

The City of Vancouver has a Citizen Dashboard that allows residents to track the progress of city initiatives. A dashboard self-proclaimed to be displayed in various layouts (charts and maps), several statistics regarding the number of projects completed and currently underway, and the average completion rate [144]. The dashboard also includes a project status tracker, allowing users to see the progress of individual projects in real time.

The WeSenseIt app in Amsterdam uses crowdsourcing to collect data from local citizens on various issues, including noise pollution and water quality [179]. Users can submit photos and information on local noise sources, map them, and make them available to city officials. The app also provides real-time data on water quality in city canals and information on air quality and traffic congestion. WeSenseIt is just one example of the growing trend of cities using mobile apps to collect data from citizens.

Another example is SeeClickFix [240], which allows users to report non-emergency issues such as potholes and graffiti, and MyNoise, which allows users to share noise complaints with their neighbours. As cities become more dependent on data to improve services and make decisions, these apps could play an increasingly important role in helping officials understand the needs of their constituents.

Smart city hackathons have gained significant popularity as a means of actively engaging citizens in solving various challenges faced in urban environments, ranging from traffic congestion to energy efficiency. These events serve as platforms for the development of innovative solutions, including applications designed for city employees, such as mobile applications designed to streamline street cleaning operations.

In the Netherlands, Amsterdam organises an engaging hackathon that brings together researchers from diverse fields to tackle pressing issues such as air pollution and traffic congestion. Similarly, the UK's Smart Cities Challenge, held in May 2018, was a collaborative effort between IBM, TechUK and the Department of Digital, Culture, Media & Sport (DCMS). The primary focus of this hackathon was to leverage data and technology to improve urban living standards.

2.3 Exploring various approaches to Ontology representation

Ontologies, as a crucial component of semantic technology, have the capacity to endow systems with the capability to model complex concepts [137]. This attribute can significantly impact User Experience (UX) by simplifying and improving the interaction between users and systems. This sub-chapter will elucidate the argument that integrating ontologies can improve UX, starting from the fundamental notion that ontologies empower systems with the ability to model intricate concepts.

An ontology can be understood as a formal representation of knowledge that includes definitions of concepts and their interrelationships within a specific domain [57]. In essence, ontologies serve as a 'knowledge map' that can elucidate the intricate relationships between diverse concepts.

Ontologies improve UX by promoting a better understanding of the data structure of the system, facilitating more efficient and effective interaction [265]. Ontologies allow the user to navigate the system more easily by providing a clear and structured representation of the data, thus reducing cognitive load and enhancing the usability of the system [298].

They foster semantic interoperability, which can significantly augment the UX. By defining a common vocabulary and establishing shared semantics, ontologies facilitate data integration and exchange between various systems and platforms [328]. This attribute can provide users with a consistent and unified experience across multiple services, contributing to a seamless user experience.

Moreover, ontologies enable personalisation and adaptivity, the cornerstones of contemporary UX design [338]. By understanding the relationships and preferences of the user within a given context, ontologies can help tailor services and information delivery according to the user's unique needs and preferences. This level of personalisation can significantly enhance UX by creating a more engaging and user-friendly environment.

While exploring various modalities of enhancing User Experience through semantics in the context of energy performance web applications and other information systems solutions, it becomes essential to address the pivotal role data ontologies play. In information science, ontologies serve as an instrument for knowledge representation, encompassing concepts within a domain and their relationships [136]. A semantic representation of these data ontologies provides an impetus to the intelligibility and efficiency of information systems.

The primary objective of this chapter is to explore various approaches to the semantic representation of data ontologies. This exploration is integral to understanding how these methods can be applied in the context of energy performance web applications and other information systems solutions.

Semantic representation approaches, ranging from descriptive logics and RDF schemas to OWL and topic maps, equip these systems with the ability to model complex concepts, relationships, and constraints, allowing more accurate, efficient, and effective data management [17]. These representations are particularly relevant for energy performance applications, where data complexity, diversity, and dynamics require sophisticated models to capture and understand various energy use and performance patterns [390].

In energy performance web applications, the structured representation of data ontologies can facilitate the interpretation and application of energy consumption and production data, leading to an improved User Experience. In parallel, in broader information systems solutions, semantic ontologies provide a structured way of representing knowledge that enables systems to understand and respond to user needs and preferences with more accuracy [177].

Therefore, exploring the semantic representation of data ontologies lies at the intersection of improving user experiences and harnessing the power of complex, diverse, and dynamic data within energy performance web applications and other information systems solutions. Through this chapter, the researcher aims to unpack the complexities and potential of these approaches, thereby contributing to the overarching objective of enhancing User Experience through the power of semantics.

2.3.1 Implementing a Digital Twin Ontology Model

The built environment is undergoing a metamorphosis due to advances in innovative technology. Wireless devices and sophisticated software are reshaping the energy and water transportation management paradigms.

However, merely facilitating communication between these technologies does not guarantee a successful interaction. Therefore, a more sophisticated solution involves leveraging the Internet of Things to bridge these resources, mainly because effective communication is hampered if diverse terminologies and definitions are used. Here, semantics plays a vital role. Using standardised semantic models, resources can build a framework for semantic alignment, enabling cooperation and the creation of advanced applications supported by a robust weather system. Semantic models define the vocabulary, information, and logic of a domain.

Historically, ontologies have been implemented to foster semantic interoperability between web services. A particularly relevant example of this within CUSP involves the various data science modules, where the overhead of term definition and labelling has been significantly reduced. Ontologies also facilitated the platform to comprehend queries and reason with knowledge without undermining performance. These data design decisions, ingrained within the design of the underlying ontology, fall outside the primary concern of this chapter.

Professional users of a Digital Twin monitoring platform, typically with backgrounds in architectural, structural, civil engineering and energy efficiency, are not expected to learn computer programming similar to industrial applications. Instead, they can apply a structured methodology for designing an interactive Building Information Modelling (BIM) model within a web application context, which includes preparation and planning, co-creation, technical facilitation, and prototyping. A collaborative interface empowers users and active citizens to design and incorporate rich content into BIM web applications. Facilitation is of great importance at this juncture in the development of ubiquitous web applications and semantic back-ends and frameworks.

The proposed Digital Twin (DT) ontology encompasses the definition of concepts and their relationships relevant to specific DT applications, including assertions and domain rules, cardinality, and the semantics of conceptual relationships in specific domains. Consequently, the proposed DT ontology model encapsulates all conceptual information within the DT domain. Figure 1 visually represents an asset's general DT ontology model during its operating lifecycle phase. The system architecture for intelligent and predictive asset maintenance is a model for this ontology model.

For simulations and analyses of digital twins (DT) related to assets, an ontology model serves a dual function. First, it collects domain information and maintains the semantic coherence of the asset functions. Second, it inherits the core attributes of a DT asset. The ontology model presented here embodies the operational phase and enables asset behaviour analysis by offering an operational understanding of crucial semantic concepts. Thus, a DT can serve as a representation of an asset currently in use.

The open architecture of the DT is structured by segmenting the model's general semantics into three preordained strata of information flow: the physical layer, the data layer, and the model layer. Specific classes have been designated for each layer of the model. In addition, an ontology model has been conceived for future databases, which presents an alternative to current industrial practices of managing large-scale data and databases.

The process of modelling a digital twin involves a structured sequence of steps, each contributing significantly to the development of a comprehensive digital representation. Initially, the procedure begins with mapping, where the proposed ontology model classes are meticulously aligned across the different functional layers of the method or process that is used for analysing asset behaviour. This alignment is crucial to ensuring that each aspect of the asset's behaviour is adequately represented in the model.

Following the mapping stage, there is the assignment task. In this phase, each class within the proposed ontology model is assigned its specific set of key data components. In addition, the types of these components are precisely defined, establishing a clear framework for the data structure within the model.

The next step involves conversion. During this phase, the relations between classes are transformed into object properties. Subsequently, the corresponding data components are integrated as data properties, complete with logical limits. This conversion is essential to create a valid and functional ontology model.

After conversion, the model undergoes a transformation process. The ontology model is transformed into a relational data structure. This transformation is facilitated by the use of cardinality and keys, which are instrumental in defining relationships and hierarchies within the data structure.

Finally, the process concludes with the populating. In this stage, real-world datasets are used to populate the relational database model. This use of actual data is vital for ensuring that the digital twin accurately reflects the real-world asset, thereby enhancing the model's reliability and applicability in practical scenarios.

In a Digital Twin lifecycle context, selecting an appropriate transformation and mapping approach is vital for an accurate and fair digital representation of real-world assets. Consequently, resources

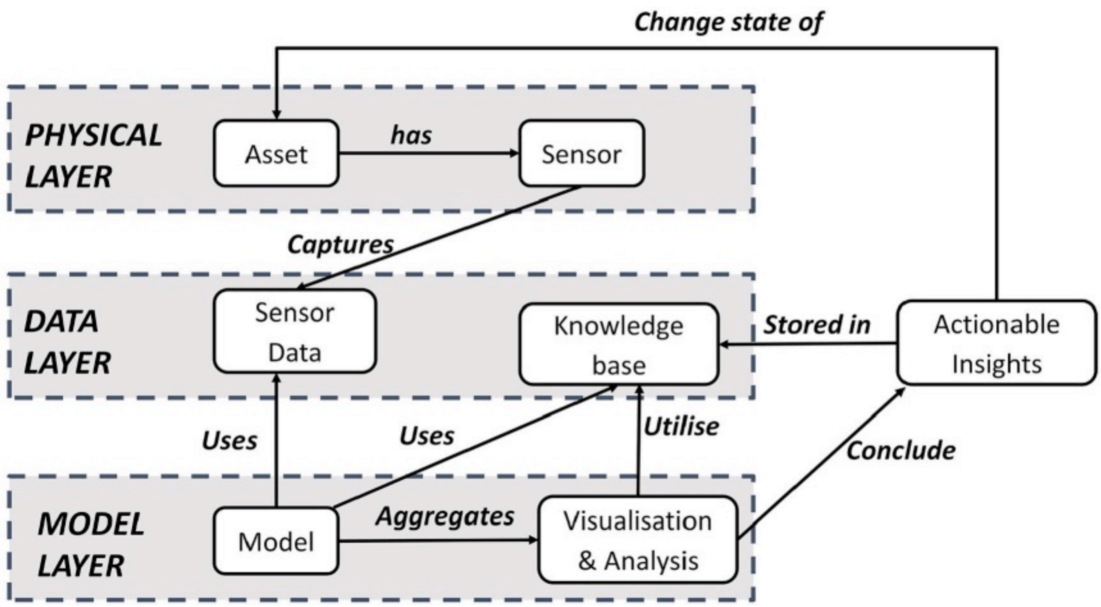


Figure 1: A Digital Twin ontology modelling approach

should be allocated to identify the most efficient technical representation while considering an ontology-based approach for the data model.

The Semantic Web of Things enables machine-readable ontologies to be accessible to web agents. When a web service is supplemented with markup, each term elucidated by the markup is linked to a specific class within the associated ontology. This system promotes the interconnectivity and usability of data throughout the digital sphere, enhancing the accuracy and efficiency of data interpretation and use.

2.3.2 HTML / XML vs OWL

From a technical perspective, context can be conferred in many ways on content or real-world entities in the virtual environment. Two of the most widely adopted representation formats are markup languages (HTML, XML) and ontology representation languages (OWL). HyperText Markup Language (HTML) (HyperText Markup Language) communicates the layout and format for the information displayed (the mode and the layout) to a Web browser.

Therefore, it is a language to format web content. However, since it is only a markup language, it lacks the semantic complexities in its syntax of markup tags to consider it a semantic language. Its successor, XML (Extensible Markup Language), is the "assembly language" of the Web that represents the significance of data for machines, injecting semantics into web contents through element tagging. However, XML is still insufficient for drawing semantics, as it does not enable users and application architects to assign meanings to contexts through semantic fields despite enabling syntax definition.

2.3.3 Graph Query Languages vs Relational Databases

Although markup and ontology representation languages can represent data meaningfully and provide context through semantics, they are relatively static representations with limited scalability. Databases and Graph Query Languages provide the ability to store indexed data and retrieve records of interest rapidly and efficiently without performing a sequential search across the entire spectrum of storage.

Regular relational databases are less semantically complex than ontologies and are represented through relational schemas rather than entity relationships. The purpose of database schemas is to represent the logical structure of the data and to build interactions between users and the domain requirements of the data. In relational databases, the procedure "JOIN" combines relations between different elements of the table. To perform this task, index lookup operations and table-to-table comparisons are required. On the other hand, the edge connecting two entities in a graph does not necessitate an expensive "JOIN" operation because graphs record entities and their interactions as nodes and edges. Semantic-based query languages, such as GraphQL (Graph Query Language), feature several serious flaws in this context. Furthermore, the use of graphs in real-world applications and with legacy Information Technology (IT) systems has its own set of challenges.

Utilising graph language commands such as Cypher and SPARQL for querying within a graph database environment is characterised by a considerable level of intricacy. This complexity poses significant challenges, particularly for novice programmers and users who are unfamiliar with these systems. One of the primary issues is the lack of uniformity in the syntax used to query in different languages. This lack of standardisation exacerbates the difficulty, as it necessitates the acquisition of specialised knowledge pertinent to each unique query language syntax.

Moreover, language-specific graph databases often have their own APIs, further amplifying technical barriers. These barriers are mainly due to the extensive and specialised technical knowledge required to effectively use these systems. Consequently, it would be immensely advantageous if users could directly query the system without the prerequisite of mastering the complexities of novel query language syntaxes. This approach would significantly streamline the process and make it more accessible to a broader range of users.

For transversal and insertion languages, the databases that Graph Query Languages could query lack standardisation and uniformity, leading to a vast range of implementations and approaches to frameworks for data interaction. Consequently, due to a lack of consistency, the researcher should dive deep into several widely accepted implementations before deciding on the most appropriate solution tailored to the particular use case.

The choice between graph-based and relational databases shall be made on the basis of the requirements of the target system. Even if they represent an integral part of the ontology-based approach towards data management of a digital twin, this subsection does not cover graph-based databases. The reason behind this rationale is that the research presently applied is primarily focused on providing a solution to support multi-user environments and a standardised language to query data and information to each user. Graph-based data models are essential for systems that need dynamic data modelling of highly connected and complicated data. Research in GQL can be employed for further study when the consistency of information retrieval from multidomain DTs is examined.

In particular, relational databases and graph query languages have plenty of built-in support for multitasking and concurrent connections between multiple users. On the other hand, many graph-based techniques do not enable multi-user environments. For example, in Neo4j, all user management is done at the application level due to its query language. Vicknair et al. [366] and Sameshima and Kirstein [318] dive into security in more depth in their articles.

2.3.4 Digital Twin Ontology Modelling Approaches

Semantic models are very often modelled in OWL (Ontology Web Language), an open data format without proprietary data involved. This feature allows all other application components to read the defined properties. The main strength of OWL is that there are no proprietary data involved. Furthermore, the OWL is much lighter (in size and computational power required) than other semantic formats, such as IFC. In addition to HTML, OWL would precisely indicate the category and real-world representation of the data (a street segment, an interior/exterior building wall). The core of the process is the ontology-to-conceptual data model. El-Ghalayini et al. [114] proposed mapping rules as guidance in transitioning from domain ontology to conceptual data model schema (CDM).

Vysniauskas et al. [370] study a different approach to transform a domain ontology into a relational database based on OWL2DB algorithms. An OWL document is analysed for the Data Definition Language (DDL) scripts. The system first converts ontology classes into data table definitions throughout this analysis and translation. In the final step, the database is filled with class instances once the objects, data type properties, and constraints are entered into a Data Definition Language (DDL) statement.

As a starting point, the technique employs a breadth-first search to identify all classes and subclasses in an ontology's hierarchical levels. The first-broad search is used to convert OWL properties into table relations. One-to-many or many-to-many relationships between tables can be constructed on the basis of the defined cardinality of class characteristics.

2.3.5 Data management needs in a Digital Twin context and uncertainties of their future

The complexity of managing comprehensive data within the context of digital twins (DT) becomes apparent in the replica of real-world entities. As industries move toward service-orientated approaches, datasets are increasingly crucial in defining product boundaries and driving innovation.

However, data-related challenges arise in managing DTs throughout various stages of development, from design to asset management and maintenance. The diversity of data types generated during the product lifecycle, including simulation, design, and manufacturing data, poses challenges in integrating and consolidating data from different sources into a unified platform such as a Digital Twin. Furthermore, data mining techniques and their applicability to DTs require further exploration and development.

Furthermore, the convergence of big data within DTs presents significant governance and storage-related challenges. Although data lakes have emerged as a potential solution to managing big data, there is still a lack of sufficient governance and technological integration. Addressing these data-related issues is crucial for fully harnessing the potential of Digital Twins in transforming industries.

Semantics, as a foundational aspect of information processing, holds promise in resolving some of the challenges associated with data management in the context of Digital Twins. By incorporating semantic technologies and ontologies, which provide a standardised and machine-understandable representation of data and its relationships, interoperability and integration of various data sources can be facilitated [173]. Semantics enable a common vocabulary and shared understanding, allowing seamless data exchange and integration between different systems and domains [41].

Through semantic technologies, data can be harmonised, linked, and contextualised, addressing issues related to data organisation, integration, and retrieval in the context of Digital Twins.

Furthermore, semantic reasoning capabilities can improve data mining and analysis, allowing deeper insight and discovery of knowledge from complex and heterogeneous data [302]. Incorporating semantics into digital twin data management processes can improve data governance, integration, and analysis, fostering a more comprehensive and effective use of data resources.

Any semantically enhanced platform has excellent capabilities to deploy artificial intelligence (AI) technology. Although various AI and machine learning techniques have been explored, this thesis focuses on the use of a simple backpropagation artificial neural network (ANN) due to its well-understood nature and suitability for research [248]. However, AI and machine learning continue to evolve, and methods such as deep learning, random forest, and ensemble learning are expected to improve prediction performance in the future [383]. Although a comprehensive evaluation of these techniques is beyond the scope of this thesis, their potential is worth noting.

Semantic platforms can greatly aid unsupervised machine learning by providing a large and well-labelled data repository. Although this thesis employs supervised machine learning methods, unsupervised approaches, such as clustering and dimensionality reduction, offer a more "plug-and-play" solution where insights are derived from available data without explicit guidance [171]. Furthermore, unsupervised learning techniques excel at ongoing data analysis and adaptation, addressing the challenges of data-driven modelling as systems evolve and user behaviour changes [220].

The growth of Internet of Things (IoT) devices in consumers' homes presents boundless opportunities for the application of AI soon. Integrating IoT-controlled systems, such as lighting, appliances, and heating/cooling, allows fine-grained control and breakdown of energy consumption within buildings [156]. Moreover, the deployment of smart metres of the IoT and the concept of vehicle-to-grid technology offer possibilities for dynamic pricing and load balancing services, facilitating cost reduction and grid optimisation [141].

In the future, responsible innovation requires anticipating the implications of emerging technologies for users and industries. The concept of a digital twin, which encompasses AI, IoT, big data, and robotics, represents a replica of a physical system that can operate with or without human intervention [23]. Understanding and harnessing the potential of digital twins can lead to transformative advances in various scientific and disciplinary domains.

The concept of human replication originated in the mid-20th century. It was introduced in various philosophical terms, such as "automata" [229] and "doppelgangers" [295]. Tao et al. [348] defines a digital twin as a virtual representation of a physical asset or system that is constantly updated to reflect operational changes and can predict the future of the corresponding physical counterpart based on data and information collected online. In the reviewed literature, the researcher has identified that the first time the phrase "Digital Twin" was used was in 1994 in medical imaging [300]. However, rudimentary versions of Digital Twins have been used since the early 1960s as computational models of medical devices called "phantoms", which were used to simulate reactions of the human body to external stimuli, such as radiation. Digital Twins were also used later in other scientific fields, such as aerospace engineering, to detect consistency in geometric and material details up to the microstructure level, including manufacturing anomalies [355].

More recent industry reports announce the advent of digital twins on an increasingly larger scale [125] [198]. However, a recent study that attempted to quantify the socioethical benefits and risks of using digital twins in health care determined that even professionals exhibited inconsistencies in understanding the term “Digital Twin”. Despite the general acceptance of benefits in terms of productivity and digital interaction, the innovation brought about by rich Web applications, such as Google Maps, could merit greater attention.

Although the literature on digital twins focuses heavily on technical architectures, exploring their application through a human-centric lens aligned with UX principles can pave the way for more intuitive and enriching implementations of this technology.

2.4 A User Experience primer

This subsection now delves into the core concepts and theories of UX. This provides a foundation for understanding the nuances of designing and evaluating user-centric systems.

The term “user experience” was first used by Norman in the 1980s to extend the excessively narrow scope of usability and to cover all aspects of a person’s experience with a system [267]. According to Norman and Draper [269], user experience (UX) refers to the emotions and attitudes associated with the use of a specific product or service, including interface, graphics, industrial design and physical ergonomics.

Alben [8] proposed one of the following definitions of UX: “all aspects related to the way people use an interactive product: the feel of the product in their hands, the understanding of how it works, the feeling during use, the achievement of their goals, but also its suitability with the global context in which they use it”.

Since the 2000s, the term “user experience” has been widely used, but understood in various ways [202]. However, many definitions have yet to be proposed without giving rise to an absolute consensus. If a precise and “operational” definition of UX is difficult to envision, it is because the user experience takes on all the complexity of the human experience.

In user experience (UX) design, the interaction extends beyond the simple binary relationship between a user and a system. It is a complex and multifaceted scenario that involves various elements and dimensions. This perspective is supported by the concept of “coexperience”, which views user experience as an interaction that is not limited to a single user, but includes the collective experience of multiple users [32].

A single approach or methodology does not drive UX design. Instead, it is characterised by many approaches with unique perspectives and contributions. For example, some approaches focus on using persuasive technology, while others emphasise the importance of understanding the context and environment of the user [133].

Kuniavsky [201] provides a comprehensive overview of current approaches to interaction design, including a collection of case studies and a review of relevant theories. Similarly, “A Project Guide to UX Design" outlines the various approaches to UX design projects, making it a valuable resource for practitioners and students in the field [358].

The diversity and abundance of these approaches reflect the inherent richness of UX design. They provide designers with various tools and perspectives that allow them to create more effective and engaging user experiences.

As defined in International Standards Organisation (ISO) DIS 9241-210, user experience (UX) is "a person's perceptions and responses that result from the use or anticipated use of a product, system, or service"[105, 246]. The ISO definition is more in line with the current understanding of UX [207], the aim of ISO is consistency, and industry ISO standards are recognised [44].

2.5 User Experience Components

In the field of academic research, there is a consensus among researchers and practitioners that user experience (UX) is the product of dynamic interaction between three critical elements: *the user*, *the system*, and *the context of interaction*. This conceptualisation is well articulated by Hassenzahl and Tractinsky [148] in their seminal work. They argue that two primary components shape the UX. First, it involves the internal state of the user, which encompasses a range of factors such as predispositions, expectations, needs, motivation, and mood. Second, it includes the characteristics of the system with which it is interacted. This state subcomponent encompasses aspects such as the complexity of the system, its intended goals, usability, and available functionalities.

Additionally, the context or environment in which this interaction occurs plays a crucial role. This context may be defined by the organisational or social framework within which the interaction occurs, the inherent meaning of the activity, and the user's willingness to engage with the system. This holistic view underscores the multifaceted nature of UX, highlighting that it is not just about the system's design, but also heavily influenced by the user's psychological state and the surrounding environmental factors.

In human-computer interaction (HCI) and design, user experience (UX) is a term that encapsulates a wide array of factors that affect how a person interacts with a system. Hassenzahl and Tractinsky [148] proposed a model of UX that effectively captures its multifaceted nature.

2.5.1 The User

Within the domain of user experience, the user component is critical as it mirrors the internal state of the individual and its subsequent influence on the interaction with a system. This component is multifaceted, composed of the user's predispositions, expectations, needs, motivation, and mood (Hassenzahl, 2008).

Predispositions, the user's ingrained attitudes or beliefs about the system, act as a personal framework through which the system's capabilities are interpreted [362]. These predispositions and expectations, user anticipations of the use of the system, are shaped by cultural, social, and past interaction experiences, and they set the stage for the engagement process [150].

Needs encapsulate what users seek from the system, ranging from functional to emotional gratifications. User satisfaction is often based on how these needs are addressed by the design of the system [393]. Whether intrinsic or extrinsic, motivation is vital as it drives the user toward the use of the system [312]. The mood at the point of interaction is an ephemeral but impactful element that can colour the user's experience, affecting their perceptions and interactions with the system [268].

This intricate tapestry of cognitive and affective elements sets the backdrop against which users interact with systems. Therefore, the design of the system must account for these variables, recognising the feedback loop between the user and the system where each influences and reshapes the experience of the other [393].

To illustrate, consider a scenario where the system is designed to be minimalistic and efficient. A user with high expectations for quick and streamlined processes may find that such a design aligns well with their needs, thus improving satisfaction. On the contrary, a different user might perceive the same design as impersonal or lacking, driven by differing predispositions and needs [362].

As we transition to examining the system component of Hassenzahl's model, it becomes clear that the system's design elements and functionalities are not standalone factors but interact dynamically with the user's internal state. This intersection underscores the importance of user-centred, empathetic design [277]. Systems tailored to the nuanced spectacle of user expectations, motivations, and moods can yield enriching experiences that resonate on a deeper, personal level (Norman, 2004).

In conclusion, a robust literature review of Hassenzahl's user experience model requires careful integration of the user component within the broader interplay of user-system interactions. Using insights from foundational work [150, 312, 362], we can articulate a comprehensive understanding of how deeply user internal states can affect and are affected by system design and interactions, crafting a pathway to designing enriching user experiences.

2.5.2 The System

The concept of "system" within the realm of user experience (UX) is explored extensively. The system, as a term, is broadly utilised in the UX literature to encapsulate various elements that play a crucial role in a user's interaction with technology or services. Although it does not fully capture nuanced interactions that involve simpler objects or human interactions, this term is widely adopted due to its comprehensive nature [39]. It represents the complexity of the interaction scenario, which often involves multiple components, each contributing significantly to the overall UX [194].

The complexity of the system, its primary goal, usability, and functionalities are integral elements that facilitate or hinder user interaction, thereby profoundly influencing UX. The complexity here is understood as the degree of difficulty or ease with which users can navigate the system. The goal of the system pertains to its primary purpose, while usability focuses on the user-friendliness of the system. Functionalities, on the other hand, refer to the various features that the system offers. These aspects are critical, as they collectively shape the user's interaction experience.

In a more granular sense, the "system" under investigation in UX contexts could include the physical device (e.g., laptops, smartphones), the software it runs (e.g., operating systems, browsers), the services it provides access to (e.g., social media platforms, online banking) and even the human elements involved in the interaction (e.g., customer service representatives, other users) [39]. Additionally, the supporting infrastructure, such as the user's internet connection or the servers hosting a website, is also considered part of this "system" [377].

Focusing on specific contexts like recommendation systems, the UX is shaped not only by the algorithmic efficiency, but also by the overall interaction with the system. This includes aspects such as the user interface, the relevance of recommendations, and the user perception of the system [194, 292]. Similarly, in the context of learning management systems, factors such as design, functionality, and user familiarity significantly influence UX [107].

It is important to note that users often perceive their experience with these various components of the "system" as a singular unified interaction, rather than differentiating or recognising these components separately [201]. This holistic perception is shaped by several factors, including the features of the user interface, the content of the website, and the user's experiences at different stages of their journey [392].

For example, in the domain of building information modelling (BIM) platforms, the user experience is determined not only by the quality and versatility of the renderings of 3D model visualisations, but also by the usability of the tools for navigation and markup, the availability of relevant parameter data and the overall workflow during different phases of the project lifecycle [243]. A user may perceive the model viewing capabilities in Autodesk Revit positively but experience difficulties finding the interface for editing object parameters. The experience they perceive when collaborating with other project stakeholders on the platform further influences holistic UX.

Therefore, in UX research and practice, the term "system" serves as a comprehensive lens through which all elements involved in user interaction are considered. This holistic approach enables researchers and practitioners to examine how each component of the system contributes to overall UX. This understanding is crucial in designing interactions that are not only effective and satisfying, but also pleasant to the user [207].

2.5.3 The Context

The third element, often overlooked, is the context in which the interaction occurs. The context includes the organisational or social framework within which the system is being used and the significance or meaning of the activity. It may involve physical location, time, cultural environment, and the role or task of the user. For example, using a smartphone for leisure at home is different from using it for an urgent work task on a crowded train. The context can alter users' expectations, motivations, and needs and transform their UX.

The components of UX proposed by Hassenzahl and Tractinsky [148] highlight the importance of a holistic perspective, acknowledging that UX is determined not only by the characteristics of the system, but also by a complex interplay of personal and situational aspects. Consequently, designing a good UX means considering not only the system and its usability but also the user's internal state and the context in which the interaction occurs. Their model advocates that to enhance UX, designers need to pay attention to the user's cognitive and emotional needs, the design and functionality of the system, and the environment and situation in which the system will be used.

The system includes all products, services, and infrastructures involved in the interaction when using the examined product.

The context defines environmental, social, and temporal factors, and (optionally) the context of the task for the experience. The user component refers to the person who interacts with the mental and physical states of the system.

User Experience (UX) is inherently subjective and deeply influenced by the user's emotional and psychological state. This subjectivity arises because users interact with the system based on their unique perceptions, expectations, and emotional state. The seminal work of Norman [265] on emotional design explicates how the user's emotional responses to a system shape their overall experience. According to Norman, the emotional impact of design plays a crucial role in user interaction, influencing both satisfaction and perception of the usability of the system [265].

Moreover, Hassenzahl [149] emphasises the importance of the user's state in shaping UX, advocating that UX is a momentary, primarily evaluative feeling (good-bad) while interacting with a product or service. This definition underscores the dynamic nature of UX, where the user's emotional and cognitive state at the time of interaction critically affects their experience. This concept aligns with the work of Wright and McCarthy [378], who argued that UX is a complex phenomenon influenced by the user's state, context, and system characteristics.

The feedback loop between the user's state and their perception of the system is also noteworthy. As posited by Garrett et al. [127], the quality of user experience affects their emotional state, which in turn influences how they perceive and interact with the system in future instances. This cyclical process highlights the interdependent relationship between the state of a user and their UX.

Within the realm of user experience (UX) studies, "context" is a multifaceted construct that encompasses a variety of elements and conditions that, while not integral parts of the system being used, can significantly influence the user's experience. This includes the physical environment, social dynamics, temporal factors, and the specific task the user tries to perform.

The **physical context** element includes the tangible sensory aspects of the environment in which the interaction occurs. Such conditions can be seen, felt, or heard, including physical surroundings, movement, temperature, weather conditions such as rain or humidity, lighting, current location, and ambient noise [3]. For example, a crowded location could influence UX if it physically affects the use of a mobile browser, such as the need to avoid bumping into others while walking and browsing.

The **social context** refers to the social norms, expectations, and influences of other people on the user's interaction with the system. The social context might affect a user's willingness to engage with a system in specific settings, such as using a mobile device privately at home versus publicly in a busy cafe.

The **temporal context** refers to the time available to the user to engage with the system, which can often be limited or shaped by the constraints of the current situation. For example, a user may need to quickly find information about their bus route to avoid missing their ride, creating a time-sensitive context.

The **task context** relates to the specific goals and tasks the user tries to accomplish while using the system. The nature of these tasks can influence how the user interacts with the system and what they need from it. For example, a user might interact with a search engine differently if he or she is trying to find a quick answer to a trivial question versus doing extensive research for a report.

Having elucidated the critical components of user experience (UX), the user, the system, and the context, following Hassenzahl's model [148] in the preceding subchapter, the researcher is now poised to dive into the methods for evaluating UX. It has been established that the user, their interactions with the system, and the context of these interactions are fundamental to UX. A comprehensive understanding of these components enables the researcher to discern which aspects might require attention when improving UX to build information modelling (BIM) web applications using semantics [307].

As such, understanding UX requires a holistic approach that considers the user's emotional and cognitive state. The subjectivity of UX means that designers and researchers must consider individual differences in users' responses to a system. As these theoretical frameworks suggest, the user's state is not just a component of UX; it is central to shaping their entire experience with a system.

However, the accurate measure of UX success is the theoretical understanding of its components and the practical application of this knowledge in the design and development process [207]. This realisation requires us to examine how UX can be evaluated and improved, leading to the subsequent subchapter: "User Experience Evaluation Methods".

2.6 User Experience Evaluation Methods

Building on an understanding of UX theories, we now focus on exploring the methods used to evaluate user experience. Evaluating UX is essential both to assess the usability of existing systems and to identify areas for improvement.

This section focuses on user-centric design, placing user needs, motivations, and experiences at the centre of the design process [269]. This approach is crucial to creating systems that are tailored to user requirements and preferences and to improve the overall user experience [293]. User-centric design is not only about understanding user needs, but also about involving the user in the design process, ensuring that the system is designed from the user's perspective [29].

Usability, an essential aspect of UX, refers to the degree to which a user can interact with a system efficiently, efficiently and satisfactorily [260]. It is a critical determinant of user experience, influencing user perception of the system and overall satisfaction [293]. Usability is not only about the functionality of the system, but also about how it fits in the context of the user and meets their expectations [260].

This section also emphasises the need for user participation in the design process, as it is vital to involve users in this process to understand their needs, expectations, and preferences, offering invaluable information for design decisions [319]. Participation of the user in the design process improves the usability of the system and ensures that the system is aligned with the user context and requirements [29, 200].

UCD (User Centric Design) evaluation methods and, by extension, UX evaluation methods, are changing and developing over time [320]. They may provide qualitative or quantitative data. Research methods may involve observations, interviews, and other longitudinal techniques to understand socio-technical problems [208].

Both industry and academia are interested in UX. The industry is interested in UX because it separates successful products from competitors [147], and the academic community is interested as UX progresses from usability, looking beyond pragmatics and towards hedonistic goals [288]. Roto et al. [308] has spotted differences between industry and academia in evaluation methods. Tan et al. [347] points out that industry and academia have taken two independent paths towards developing UX evaluation methods that lead to definitions and terminology that cannot be shared. Ardito et al. [21], has worked to narrow the gap between industry and academia. UX evaluations can also be incongruent with how companies work currently [308], although UX evaluations are critical elements of high-quality UX [305].

The disconnect between academic research and industry practice in user experience (UX) has received some attention recently [61, 346, 132]. In the early 2000s, there was concern that usability, while a critical aspect of interface design, did not adequately meet the needs of the industry [304, 205]. Similarly, in the present context, there is speculation that the academic literature on UX may not be entirely relevant or accessible to industry professionals [61, 346, 132].

There are numerous reasons for this disconnect. One potential issue is the complexity and variability of UX, which can be difficult to distill into actionable guidance for practitioners [30, 148]. Furthermore, more awareness or understanding of the relevance and applicability of the UX literature in industry contexts may be needed [369, 206].

UX shares many basic concepts with usability, but extends them to incorporate emotional, subjective, and temporal aspects of a user's interaction with a system [148, 307]. Although usability has traditionally focused on objective measures of interaction (such as efficiency and error rates), UX encompasses the broader subjective experience of a user in contact with technology. This includes perceptions, feelings, thoughts, and other personal and situational aspects that may occur before, during, and after use [150, 30].

In the UX design process, usability is often an essential aspect of the 'pragmatics of interaction.' This term refers to the functional and practical nature of a system for users and how it supports them in achieving their goals [30]. Thus, the objective aspects of usability are integrated within the broader and subjective UX framework [150, 30].

Users have several needs about their functional (cognitive) and emotional (affective) well-being, and this is included in the term "user experience" [148]. There is a connection between the subjective and objective aspects of human-product interactions [294], together with the internal and external aspects. Although the functional aspects of a product are related to its usability, affective properties focus on how that product makes an individual feel. The emotional aspects of user experience are not studied as extensively as the functional needs of users, as illustrated by Hassenzahl and Tractinsky [148], who stated that a complete understanding, definition and operationalisation of the user's emotional needs is still needed.

User Experience (UX) refers to a user's specific experience while interacting with a technology, product, or service. This experience is shaped by various factors, including user perceptions, emotions, and responses to system functionality, ease of use, and overall performance [148, 339]. UX design aims to optimise this experience by prioritising the user's needs, preferences, and context of use over focusing exclusively on the features of the product [30, 150].

Given the integral role of the human factor, it is appropriate to postulate that the processes and methods of UX design are deeply rooted in the User-Centred Design (UCD) approach [269, 339]. UCD promotes user participation throughout the design process, allowing a better understanding of user needs and preferences and ensuring that the resultant product or system is tailored to their expectations [265, 339].

To effectively adopt a UX-centric design approach, it is necessary to take advantage of different methodologies and understand UX in depth, including its nuances, principles, and the broader process within which it fits [307, 365]. A core principle of UX design is to involve users as early as possible in the product development cycle, thus designing "for" and "with" them [319, 53]. This early and continuous user participation can facilitate the creation of products that truly meet users' needs and provide optimal UX [30, 150].

In 1986, Norman and Draper [269] published the first book dedicated to User-Centric Design. They insist that developers focus product development on users' needs, pay particular attention to the analysis of their activity, and perform iterative evaluations of the system from the earliest stages of the design. Approximately ten years later, the user-centred design process is formalised in the international standard ISO standard 1999-13-407 [172] with the primary objective of designing usable systems. In 2010, this standard was revised to incorporate the concept of user experience. ISO 9-241-210 [105] reproduces and supplements the founding principles of UCD:

1. The design process is based on an explicit understanding of users, tasks, and environments.
2. Users are involved throughout product development, including design, guided by thematic evaluation phases.
3. The process is iterative and requires going back and forth between the different phases of the design.
4. The design team integrates multidisciplinary skills and perspectives using a participatory approach.

5. The design covers the entire user experience, thus integrating the components related to the system's performance (performance, functionalities) and those related to the user (habits, personalities, skills).

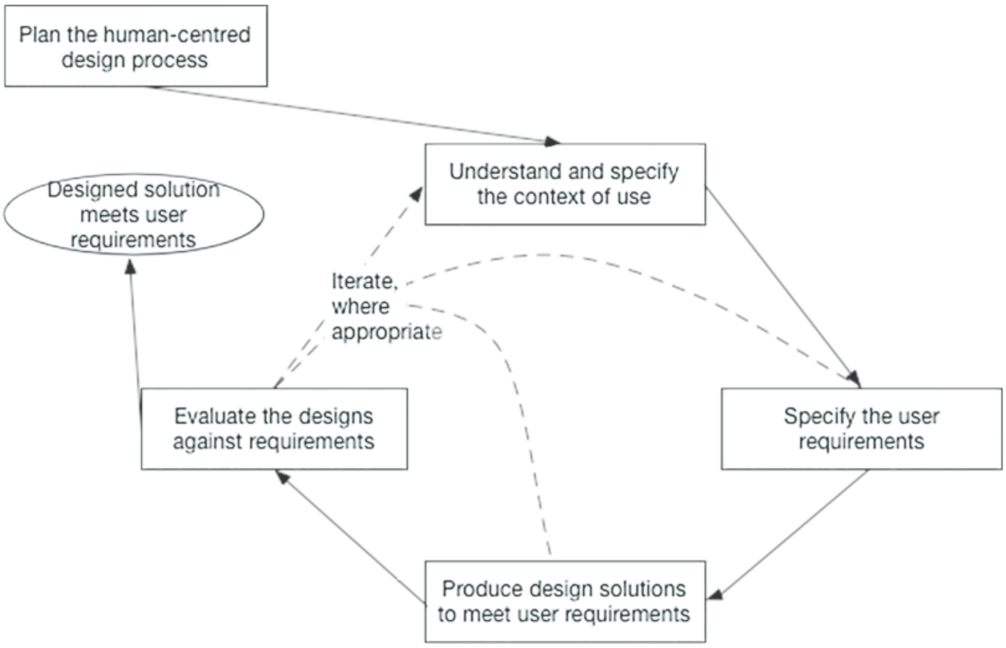


Figure 2: The User Involvement Process in the User Experience Design according to the ISO 9241-210 standard

One of the fundamental principles of UCD is to integrate the user as early as possible in the product development cycle. In other words, UX designers should design with both users in mind and together with users. Therefore, the UX design process and methods originate in the UCD process.

User Experience (UX) is a new term that replaces usability and covers a broader range of topics, from usability to experiential aspects of use [148]. Some researchers even consider usability and UX interchangeable terms [91, 365], but the researchers consider UX a broader paradigm than usability.

It is essential to remember the pragmatic aspects of interaction when designing UX. The goal is always to create systems that are easy to discover, use, and master. However, UX design focuses on pleasure, emotions, basic needs, and values. Its goal is to create products and services that attract, captivate, delight, inspire, and improve the quality of life of the end user.

To meet this challenge, UX designers use a variety of tools:

- exploration tools that help to understand in depth the needs of users
- ideation tools that stimulate creativity and support innovation

- techniques generators that guide the design process
- assessment techniques that involve the target users

The user-centred design was initially intended to ensure the usability of the systems. According to the revised version of ISO 9241-11 (1998), usability is the "degree to which identified users can use a product to achieve defined goals, with effectiveness, efficiency, and satisfaction, in a specified context of use" [52, 43].

Efficiency can be defined as "the precision and degree of completion with which the user achieves goals specified" or the "the ratio between the resources spent and the precision and degree of completion according to which the user achieves the specified objectives".

Satisfaction refers to the absence of discomfort and positive attitudes toward using the product. Some definitions of usability incorporate factors other than effectiveness, efficiency, and satisfaction. As described in [296], the ease of learning often complements the definition of ISO standards.

Having thoroughly examined the various methods for evaluating user experience (UX) in this subchapter, it becomes clear that these methods are underpinned by and intertwined with different UX theories and models. These theoretical foundations guide the construction and application of UX evaluation methods and offer valuable insight into the underlying mechanisms that drive user perceptions, attitudes, and behaviours.

Therefore, to deepen the understanding of UX, it is imperative to explore the theories and models that serve as the basis of this field. By delving into these, the researcher can better comprehend the nuances and complexities of UX, which is paramount for enhancing the UX of building information modelling (BIM) web applications using semantics.

This exploration naturally leads to the next sub-chapter, called "User Experience Theories and Models".

Although designing interfaces for complex urban systems poses many UX challenges, they underscore the growing need for human-centred research on semantics in the built environment. This leads to defining the scope of this thesis research.

2.7 User Experience Theories and Models

Having thoroughly examined the UX evaluation methods in the previous section, it becomes clear that these methods are underpinned by different UX theories and models. These theoretical foundations guide the construction and application of UX evaluation methods.

Within this sub-chapter, the researcher will examine various prevalent theories and models in the field of UX, aiming to shed light on the theoretical underpinnings of UX, which will, in turn, inform better design decisions and evaluation strategies, thus improving the usability and overall UX of BIM web applications.

2.7.1 The Hassenzahl model

The two main models used in UX research are those of Hassenzahl and Mahlke.

In the Hassenzahl model, the designer's point of view differs from that of the user (Figure 3). A designer selects and combines elements such as content, presentation, functionality, and interaction procedures [148] to instil a particular character or essence in a product. However, this choice is inherently subjective and merely embodies the designer's intentions. Therefore, assessing whether the product aligns with user perceptions and needs is crucial to foster mutual understanding between the designer and the user. This understanding is integral to the user experience design process.

From the user's perspective, the perceived quality of a system hinges primarily on two dimensions: its pragmatic and hedonic quality [148].

Pragmatic quality refers primarily to the utility and usability of the instrumental aspects of the product or system. In other words, it reflects how effectively a product facilitates the achievement of specific objectives or tasks, known as "do-goals". Characteristics of pragmatic quality include the clarity, structure, and predictability of the system [354].

On the other hand, hedonic qualities are non-instrumental and pertain to the user's self. They derive from the potential of the product to provide entertainment and fulfilment and to meet deeper human needs, often referred to as "be-goals" [151]. Aspects of hedonic quality include the ability of the system to foster a sense of achievement, facilitate connection with others, improve user control,

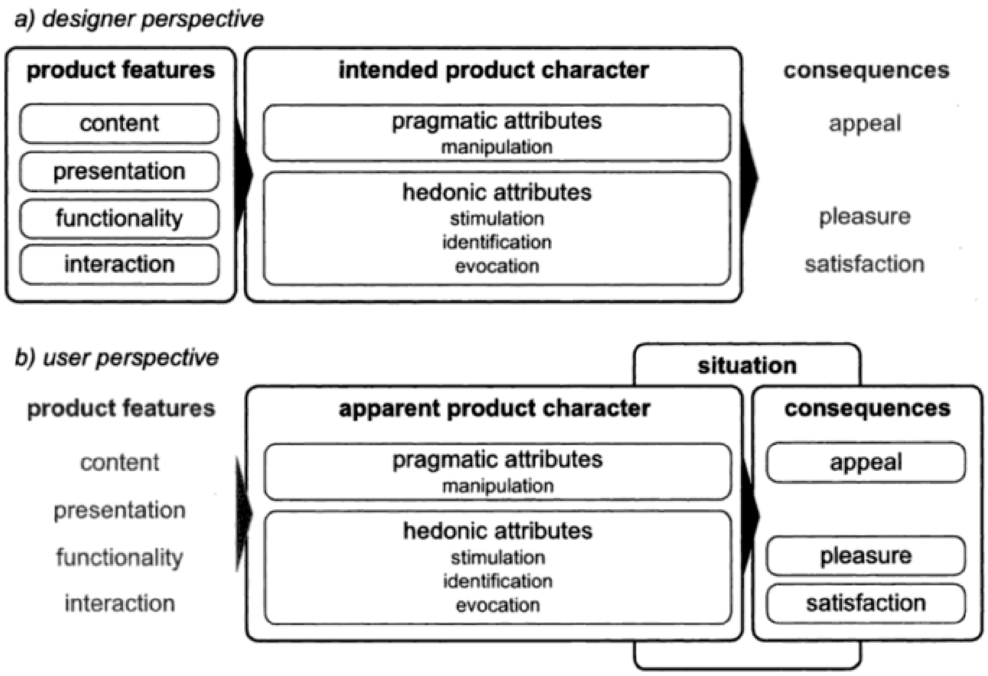


Figure 3: The User Experience Model of Hassenzahl

and confer popularity or social status.

The pragmatic and hedonic qualities that users perceive converge with an overall evaluation of attractiveness [148]. Subsequently, this assessment can generate behavioural responses, such as increased use of products, and emotional consequences, such as feelings of satisfaction or joy.

2.7.2 The Mahlke model

Mahlke's model (Fig. 4) is a more global research framework that incorporates various aspects of user experience. The three common reasons for human/technological interaction and the way users

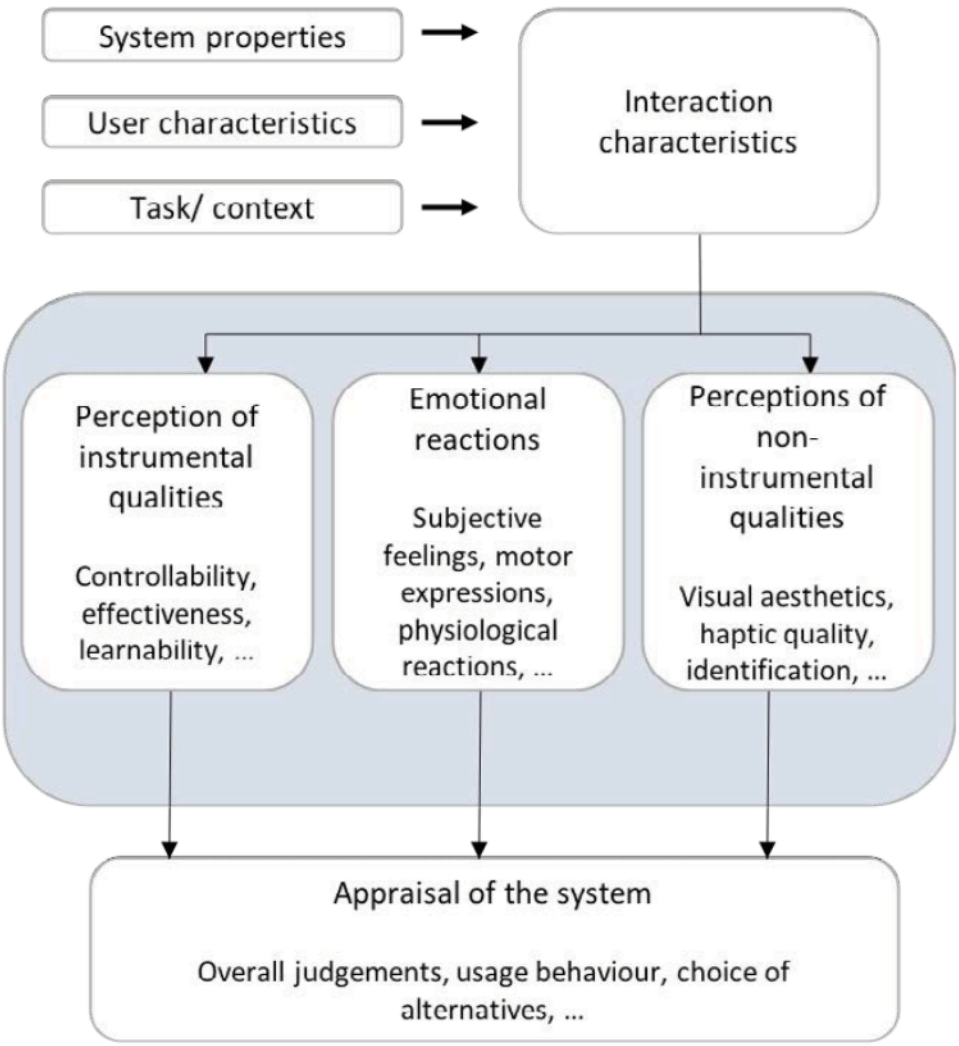


Figure 4: The User Experience Model of Mahlke

perceive interactive products are:

1. The system's properties are product-specific and related to specific design solutions. These include features, interface properties, and dialogues.

2. All the attributes of the person interacting with the system are the user's characteristics. For example, his demographic profile, expectations, needs, motivations, values, and knowledge can be mentioned.

3. The context parameters include all aspects of the system situation, including physical, social, technical, temporal, and task factors.

The UX compounds and processes are similar to those described in Hassenzahl and Tractinsky [148]: judgements and behaviours will lead to perceiving instrumental and non-instrumental qualities. The central role of emotion is increasingly emphasised because the perceived characteristics of a system and the consequences for its application play a mediating function.

Mahlke's model serves as a conceptual framework for user experience (UX) design, extending its application to several areas [350]. However, a critical aspect frequently overlooked in the discourse of UX, including in Mahlke's and other similar models, is the time factor or temporal dimension.

The temporal dimension is critical for the overall user experience, and this researcher posits that it should be acknowledged as the "fourth dimension" of UX, supplementing the triad of user, system, and context [183]. This perspective is grounded in the belief that an interaction's context is spatial and temporal.

Before diving into UX design, it is crucial to appreciate the emotional journey a user may undergo. Despite its importance, this aspect is often left on the sidelines during the UX design process [169]. It is incumbent upon both designers and nondesigners to consider the temporal dynamics of UX.

From a design point of view, considering the duration of the user experience allows designers to ensure that the characteristics of the system are maintained consistently during transitions between phases. For users, the timely availability of a product is a determining factor in its usability [266].

In essence, the lifecycle of a product encompasses several stages: it begins as a commodity (pre-use), evolves into an instrument (use), potentially becomes a habit (long-term use) and eventually may be transferred or recycled (post-use) [183]. Recognising and factoring in these stages can significantly enhance the UX design process.

2.7.3 The Karapanos model

The Karapanos model (Figure 5) divides product adoption into three phases: orientation, incorporation, and identification. Three forces will catalyse this transition (familiarity, functional dependence, and emotional attachment). At each phase, various characteristics of the qua system are observed. The orientation phase corresponds to the moment of product discovery, characterised by the stimulating effect of novelty and the desire to learn how to use the product (ambivalence between excitation and frustration).

Gradual familiarity with the product allows the user to enter the **incorporation phase**. Here, the product becomes indispensable daily, and its main characteristics are utility and long-term **usability** (beyond the **learnability** of the **orientation phase**). Functional dependence evolves into emotional attachment, where the product becomes integral to daily life and the individual's identity. Qualities that are considered valuable are those that allow identity communication and participation in social interactions.

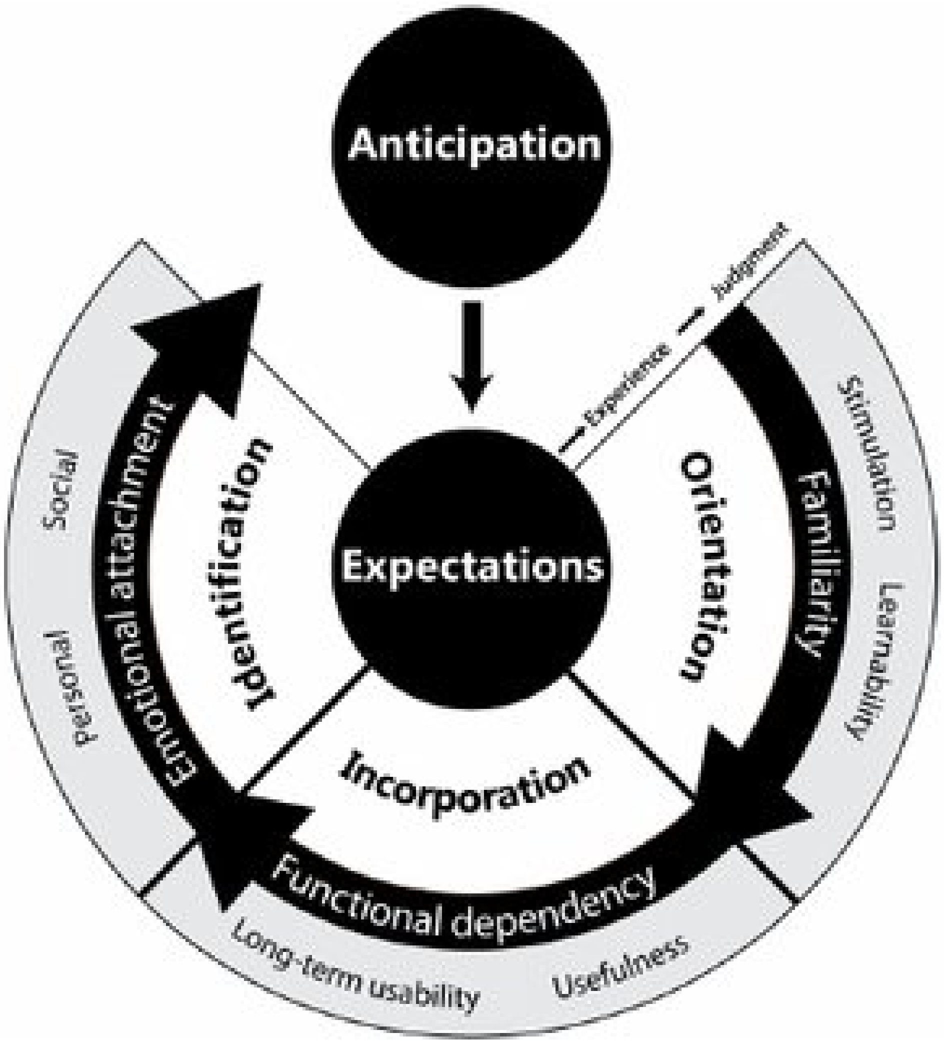


Figure 5: The temporal model of User Experience (adapted from Karapanos)

Several researchers have developed models that illustrate how the user experience evolves over long periods [183]. Using Karapanos' analytical framework, it is possible to view the user experience as a continuum that begins with the anticipation phase and progresses until a sense of attachment is formed [183]. As systems and products are increasingly intertwined with service delivery, there is a risk that users will feel a loss of control. To address this, designers must go beyond fostering initial acceptance that leads to a purchase and instead focus on ensuring sustained use of the product over time [183].

The temporal dynamics of UX has significant implications for design. Suppose that the user's experience is shaped before interacting with a product and continues for an extended duration. In that case, designers must be aware of this temporality and observe the characteristics of the system that allow for smooth transitions between different phases [183]. A user experience designer must consider the product holistically, recognising it as something that users purchase (pre-use), use (use), enjoy and adopt (long-term use) and occasionally pass on or recycle (after use) [183, 151].

While UX design expands beyond usability, it does not disregard the pragmatic aspects of interaction. The goal is to create systems that are easy to learn, use, and understand [149]. However, UX design considers pleasure, emotions, basic needs, mental well-being, and values [150, 123]. The aim is to develop products and services that attract, attract, enchant, and inspire users, improving their overall quality of life [151].

To meet this challenge and become proficient "experience designers," designers must leverage various tools. These include exploration tools that help to gain an in-depth understanding of user needs, ideation tools that stimulate creativity and support innovation, concept generation tools that guide the design process, and evaluation tools that involve target users.

Having provided an overview of the importance of user experience, the next subsection focuses on exploring the essential components that contribute to an enhanced user experience in the context of BIM web applications. This section delves into the various elements that shape the user experience landscape, examining their roles and interdependencies.

2.7.4 The Technology Acceptance Model

Although the models discussed so far have offered critical perspectives on user experience, the exploration of theoretical frameworks in technology acceptance has guided the focus of the researcher on the Technology Acceptance Model (TAM). TAM, rooted in the theory of rational action, is a key model to understand the adoption of technology. It proposes that perceived usefulness and ease of use are essential in influencing people's intentions to use and accept information systems and technology [7, 76].

TAM proposes two primary factors that influence the user's decision to use a particular technology: Perceived Usefulness (PU) and Perceived Ease of Use (PEOU).

Perceived Usefulness (PU) refers to the degree to which a person believes that using a particular system would enhance their job performance or task achievement. PU can be evaluated through user feedback, performance metrics, and comparative analysis with existing systems. For example, in a healthcare setting, PU could be assessed by how a new information system improves patient care or reduces the time spent on administrative tasks [160].

Perceived Ease of Use (PEOU): This aspect concerns the degree to which a person believes that using a system would be free of effort. In practical terms, PEOU can be measured through user interface design, training requirements, and system intuitiveness. For example, a new software application in an educational context could be evaluated based on the ease with which teachers and students can navigate and use its features [363].

In a real-world application, TAM is often employed in the process of introducing new technologies (user onboarding processes). Organisations may conduct surveys or focus groups to assess PU and PEOU among potential users. Adjustments to the technology or training provided may be made based on this feedback to improve the overall acceptance and usage of the system.

Furthermore, TAM has been adapted and extended to suit different contexts. For example, the Affective Technology Acceptance model integrates user emotional responses into the acceptance process [210], which is particularly relevant in consumer technologies where emotional involvement plays an important role.

The extensive application of TAM in various domains, such as healthcare, police, and household technology, illustrates its versatility in assessing user acceptance of new technologies [371, 160, 84]. This adaptability extends to its variants, such as the affective technology acceptance model, which considers emotional responses [210].

In the context of user experience within the built environment, especially in web applications, TAM's direct applications are less evident in the existing literature. However, its established efficacy in diverse fields suggests its potential applicability in this area. The broad spectrum of TAM's applications, from health care to disaster management, indicates its capacity to offer insights into user acceptance in varied contexts, including web applications in the built environment [160, 84].

The literature on user experience underscores the subjective and complex nature of this field, highlighting psychological factors and the importance of secondary user experiences in design [297, 42, 12]. These considerations reinforce the need for models such as TAM that can navigate the intricacies of user experience in technology.

Although the application of TAM in the context of web applications within the built environment has not been extensively explored, its proven effectiveness in various domains, including the World Wide Web and information technology, positions it as a potentially valuable model for investigating web application acceptance in the built environment [209, 249, 4, 185, 15, 20]. The focus of TAM on perceived ease of use and usefulness, validated in web contexts, supports its potential to predict actual use of technology and understand user behaviour and adoption [363, 356, 371].

Although TAM provides a valuable perspective on user acceptance, some limitations have been noted that warrant consideration. Criticisms include its parsimonious nature and the dependence on self-reported usage data rather than actual system logs [356]. Furthermore, external variables that influence perceived usefulness and ease of use may require a deeper examination [213]. Cultural differences in technology adoption also merit exploration when applying TAM in different contexts [11].

To mitigate these limitations, researchers suggest integrating objective system data, qualitative feedback, and longitudinal studies into TAM-based evaluations [356].

Complementary models, such as the fit of task technology, could be combined with TAM to assess additional adoption drivers [193]. Extensions of TAM also provide avenues for responding to concerns. The literature highlights that exploring complementary models such as the Unified Theory of Acceptance and Use of Technology (UTAUT) could provide a more comprehensive picture of the interpersonal and social dynamics that influence adoption [364].

When applying TAM to evaluate the user experience of semantic web applications in the built environment, triangulating self-reported data with behavioural metrics could prove insightful. For example, perceived usefulness could be compared with usage patterns. Including qualitative contextual data through interviews and ethnographic observations can reveal external factors that affect perceived ease of use and usefulness.

Longitudinal studies mapping acceptance over time could illuminate evolving needs and challenges in adopting semantically enriched systems. Comparative analyses of user groups with varying domain expertise can highlight differences in acceptance related to background knowledge.

The extensive use of TAM in various technological contexts, including its adaptability to web-based technologies, makes it a promising framework for exploring the acceptance of web applications in the built environment. Therefore, the literature supports the consideration of TAM in this context, highlighting its potential to deepen our understanding of user acceptance and experience in web applications within the built environment.

2.8 The Role of BIM in the general context of UX

Having examined the general role of semantics in user experience (UX), the researcher focused on a specific application of semantics within the UX domain: building information modelling (BIM). By leveraging semantic representations and digital information models, BIM enables a more comprehensive understanding and communication of complex architectural and engineering data, fostering collaboration, efficiency, and improved decision making throughout the project lifecycle [331]. This subchapter explores the integration of BIM principles and methodologies into the broader context of UX, shedding light on how seamless interaction between BIM and UX can lead to more user-centric and practical design processes in the built environment.

According to Vanlande et al. [360], BIM is the process of generating, storing, managing, exchanging and sharing building information in a reusable and interoperable manner. Azhar et al. [25] expands on this by highlighting the need for a computer-generated model to simulate the various phases of a project, from planning and design to construction and operation. Eadie et al. [111] elucidate the perspective of the UK government, which recognises the multitude of efficiencies and benefits that BIM brings across the project life cycle.

Implementing BIM enables designers to create three-dimensional digital models that include data related to the physical and functional characteristics of a building. This capability allows stakeholders, such as architects, engineers, and contractors, to collaborate on coordinated models, which improves their understanding of how individual contributions fit within the overall project. As a result, it facilitates a more efficient workflow. The dynamic nature of BIM models ensures that alterations in an element are reflected in all views, maintaining consistency and accuracy throughout the project documentation.

The strategic use of information within the models, for example, to assess lighting, energy consumption, and structural integrity, can significantly improve the quality of the design before construction begins. This preemptive approach expedites approval processes through realistic visualisations and maintains the model's intelligence from concept through to construction. During the construction phase, BIM provides valuable information that improves both efficiency and effectiveness, thus fostering a deeper understanding of future operations and maintenance requirements.

For owners, the versatility of BIM extends into aspects such as predictive maintenance, asset tracking, and facility management, as well as being an invaluable resource for planning future renovations or deconstruction. By embracing BIM, stakeholders can anticipate reduced project risk, improved timelines and cost efficiency, and better project results. The growth of BIM is further accelerated by advances in cloud-connected technologies, enabling project teams to collaborate in unprecedented ways. As the Architecture, Engineering and Construction (AEC) industry faces transformative global trends, adopting robust BIM solutions becomes essential for those who want to secure more work, deliver projects more effectively and design superior buildings.

The adoption rate of BIM varies globally, with the UK leading the way. A striking 90% of UK construction companies have integrated BIM into their practices, a stark contrast to the adoption rate of 50% in the US. This disparity underscores the UK's commitment to advancing construction industry standards and practices.

Implementing BIM processes requires careful consideration of data security, particularly when sensitive information is involved. The construction team must ensure data confidentiality, both internal and external to the BIM model. When the model is part of an extranet, it can raise legal challenges that must be addressed prioritising construction contracts to mitigate significant risks [79, 357].

In an attempt to make future predictions, the global construction market is poised for substantial growth, with an expected increase of more than 70% by 2025 [113]. The UK has set ambitious goals to drive progress, including substantial reductions in construction and lifecycle costs, project delivery times, and greenhouse gas emissions. These goals reflect a broader commitment to sustainability and efficiency within the industry.

Historically characterised by fragmentation and adversarial relationships, compounded by linear workflows, the construction industry is now at the cusp of a cultural shift. BIM emerges as a pivotal force in driving this transformation, offering opportunities to reduce energy demand, streamline processes, and reduce carbon emissions. Furthermore, Beach et al. [34] suggest that BIM fosters efficient energy modelling and interdisciplinary collaboration, contributing to a more integrated lifecycle and supply chain management.

Beyond the technological realm, BIM provides many sociotechnical benefits that enhance the process and products of architectural design. These benefits profoundly alter the relational dynamics associated with construction. As the industry grapples with stringent legislation on energy efficiency, economic pressures, competitive forces, and changes in work culture, BIM stands as a beacon of innovation and transformation.

Integrating BIM within the construction industry signifies more than technological advancement; it represents a paradigm shift in the way buildings are designed, constructed, and maintained. As the industry evolves, the role of BIM will continue to expand, paving the way for a more efficient, sustainable and collaborative future.

In this context, BIM can be a game changer force that would help move to more efficient and cost-effective practices [58][284].

The transformative role of building information modelling (BIM) and information and communication technologies (ICT) in the construction industry is increasingly recognised, particularly in improving decision-making efficiency and construction follow-up [24]. In its initial stages, BIM has demonstrated its potential to provide advanced design support through 3D visualisation, physical simulation, and upstream evaluation of design options [24]. Furthermore, it has been instrumental in the planning and continuous monitoring of construction, facilitating shorter timescales for the completion of work [24].

Integrating digital technologies into the construction process can significantly improve the energy efficiency of buildings [263], which is particularly relevant in European companies, where time and cost savings are crucial to maintaining competitiveness. The use of BIM and ICT can lead to substantial changes in these areas, contributing to the general sustainability and efficiency of the construction industry [24, 263].

In the construction industry, the integration of building information modelling (BIM) plays a pivotal role in enhancing the efficiency and sustainability of building projects. Throughout the various stages of construction, it is imperative to enrich the building information model with extensive data. These data, mainly related to building materials and both simulated and actual usage, are indispensable to support energy analyses and simulations [263]. The use of a data-driven approach, underpinned by BIM and Information and Communication Technology (ICT), facilitates a more thorough and precise evaluation of a building's energy performance. This, in turn, aligns with the broader objectives of energy efficiency and sustainability within the construction industry [263].

Building knowledge modelling, a key aspect of construction projects, significantly enhances collaboration among various stakeholders, including contractors, designers, suppliers, and facility managers. This collaboration spans a variety of design and construction activities. BIM has been recognised in academic research as an effective tool to address several critical challenges in construction projects [253, 63]. One of such challenges includes the need to mitigate project failure by enhancing coordination among project teams across supply chains, thus fostering successful project execution.

Another challenge involves navigating the complexities introduced by new procurement methods, such as "design-build-operate" contracts, which demand a more integrated approach to project management [92]. Furthermore, the use of BIM helps reduce the total cost of the lifecycle of buildings through the application of BIM in facility management, thus achieving cost efficiency throughout the life of the building [36].

BIM facilitates the systematic collection and storage of information in a BIM-compliant database. This database becomes a crucial resource for a wide range of activities, including energy management, maintenance and repair, and room management [73]. The comprehensive nature of this database ensures that all relevant information is readily available, thereby streamlining processes and enhancing the overall efficiency of building management.

Integration of BIM in construction projects not only optimises the design and construction phases, but also extends its benefits to post-construction activities. By fostering collaboration, improving efficiency, and reducing costs, BIM emerges as a cornerstone technology in the pursuit of sustainable and efficient construction practices.

2.8.1 BIM for Energy Efficiency

There are many research efforts in building information modelling and energy efficiency to develop a technique for using BIM to minimise energy consumption and pollution in buildings. Implementing BIM for energy efficiency would provide energy savings through accurate energy tracking, real-time decision support systems and actuators, and detection of consumption patterns.

Furthermore, the following can inform the optimal management of the evolution of energy use in building construction:

1. focusing on a functional approach (i.e. BIM, real-time data analysis, behaviour modelling)
2. improved monitoring of energy flows and usage in buildings
3. new collaborations between energy managers, energy manufacturers, energy equipment suppliers, and technology (including smart software tools)
4. it enables advanced analytics via KPIs (Key Performance Indicators) monitoring, annual consumption forecast improvement, reports and custom notifications [35]

Smart implementation of (reduced) building energy use would result in economic benefits in proportion to the planned reduction in energy consumption. Although, according to some thermal regulations, the energy consumption of a building is not allowed to exceed a specified cap, the actual energy efficiency is typically lower. One way to reduce the difference between prediction and reality is to improve the entire process from the initial design phase to the operational phase [285] [388].

Building energy efficiency can be further improved by implementing semantic web principles. Interoperability issues can be alleviated using semantics and ontologies in older buildings, as the building's various systems communicate using various protocols. BIM and IFC ontologies help define building concepts and allow external factors such as energy prices, weather forecasts, energy simulations, and Building Management Systems (BMS) data to be linked. Data and objects are given context and meaning at the heart of semantic web technologies. Following this approach, controllers can make better decisions for the user, making them more aware of their surroundings.

2.8.2 BIM and semantics

The construction industry and the built environment have experienced sequential growth. An analogy can be established with the gradual development of the Web environment, which started with Web 1.0, where the pages consisted of a simple HTML markup. The Web then moved on to the second generation (web 2.0), which brought the use of the semantic web and ontologies to life. Subsequently, the Web moved to the third generation, introducing and merging the user experience, IoT and *research artefacts*. It is vital to provide an overview of the semantic web as a context for the transition that BIM technologies need to undergo to discuss progress toward a web-enabled environment. The semantic web augments existing web technologies with abstraction layers that enable machines to comprehend the meaning and context of content and data. Advanced applications are enabled by introducing this abstraction layer above the data, and application development is simplified by decoupling data collection and representation from the development of logic-based software.

Alesso and Smith [10] defined The Semantic Web as "an extension of the current Web in which information is given a well-defined meaning, which better allows computers and people to work in cooperation". Similarly, ontologies can underpin BIM in a digital twin, where the user interacts with the digital replicas of actual components, such as sensor data, and provides real accounts of the performance of the artefact of interest. Extra layers of complexity are added to a building or city-level representation (i.e., specifying what objects each floor contains).

An ontology is derived from the philosophical study of "being" and is used by computer scientists to refer to computing systems capable of self-reflection. An ontology is a conceptual framework that encompasses the technological "nature of existence" of an object in the world. For example, the ontology provides a control system with an understanding of the fundamental physics of operation, how that operation interacts with the world around it, how its internal physics is organised, how the object interfaces with controls, the physical parameters of operation, and the operation "meta-semantics". This term refers to the ontology of an operation as it is translated into a language or protocol.

Abanda et al. [2] have developed ontologies related to concepts and objects of the built environment with a specific focus on the use of BIM, according to a review of the current literature. The semantic web concept ontology aims to develop a common vocabulary to describe concepts and objects through a shared model. In order to construct knowledge (rather than mere information), these texts describe the assets they represent within the real-world and the relationships between them, including rules and logic.

Kofler et al. [195] developed an OWL ontology to capture and link core concepts within the smart home to provide a stable platform for the development of smart homes. MAS. Users, processes, exterior information, energy, and resources were included in this category. With this ontology, any control platform built on top of it could make intelligent decisions for users who want to reduce their carbon footprint. The MAS control procedure developed by Ruta et al. [311] is also semantically linked. There is a central agent who represents all building devices. For overconstrained problems, the control procedure must juggle various constraints of varying importance and relax the least essential constraint.

The semantic web community has helped develop standards and technologies that facilitate human-machine intelligence interaction. The semantic web is well positioned to address the research and development gap in the interoperability of the application layer of the built environment [145, 330]. This is particularly relevant in the context of BIM, given the ongoing development and standardisation of ifcOWL.

In the last ten years, research has progressed more frequently toward a web-based version of IFC with the use, in almost all desired syntax, of a web ontology language [279, 278], which is more adaptive to use on web-based systems. This objective was to combine BIM with the data web using a consensus approach, partly by the W3C Linked Building Data Community Group.

According to Curry et al. [90], BIM models can "serve as an information backbone" for more intelligent management platforms. BuildingSMART's Industrial Foundation Class (IFC) schema is the most commonly used digital BIM representation at the building level. The cohesion around an ontology for describing the data in an IFC model towards an ifcOWL representation has continued to grow. Using ifcOWL model data and BMS sensor data, the author developed a cloud-based management platform that provides users with information on their energy consumption. [99] also used an Autodesk Revit BIM model as the central model from which more specialised software could produce modules specific to thermal and lighting. These models are web-enabled 'ontologies', an evolution of traditional entity-relationship models better suited to the web's adaptable and distributed nature. Their primary characteristics are as follows:

- they are highly extensible and distributed;
- each concept is uniquely identified; and
- they result in a network-based or graph-based model of the concepts and objects in a system.

The semantic definition provided by ifcOWL through BIM could provide the foundation for highly improved building life cycle management. This representation format can link BIM, infrastructure, and energy data.

The KnoholEM project exemplifies how semantics can be used to control and optimise a building in real-time. The concepts of energy consumption, management, and detection were added to an existing ontology based on the IFCs defined within this project. The ontology enforces consistency between BIM, energy model, and BMS objects, sharing common descriptions and being interpreted as representations of the same physical object. Based on energy simulation data and historical data mining, the semantic base assisted in generating energy savings rules through surrogate ANN machine learning models. Based on the facility manager's request for energy savings, an Artificial Intelligence (AI) developed a set of optimal rules. A simple user interface for a fuzzy inference system was used to suggest the most appropriate rules to the facility manager. For the facility manager, semantic mapping provided a 3D model of their building with the ability to select individual zones to retrieve a breakdown of energy consumption and comfort levels for each area.

Yuce and Rezgui [388] presents a system that uses semantics to address the disparity between predicted and actual energy consumption in buildings. This study introduces an integrated approach that combines artificial neural networks (ANN), genetic algorithms (GA), and rule-based semantic techniques.

The primary objective of the proposed system is to improve the precision of the predictions of energy consumption in buildings. To achieve this, the system incorporates rules-based semantic mechanisms that use the semantics of building elements and energy usage patterns. The proposed system uses these semantics to bridge the gap between predicted and actual energy consumption.

The combination of ANN and GA enables the system to optimise the prediction models by learning from historical energy data and adapting to changing conditions. Semantic rule-based techniques enhance the interpretability and explainability of the system's predictions. This allows building stakeholders to gain insight into the factors influencing energy consumption and to make informed decisions about energy optimisation.

By reducing the gap between predicted and actual energy consumption, the proposed system improves the efficiency and sustainability of buildings. Using semantics in the system improves the understanding of energy usage patterns and enables more accurate predictions. This, in turn, facilitates effective decision-making and the implementation of energy-saving strategies.

Therefore, this paper introduces an ANN-GA semantic rule-based system designed to minimise the disparity between predicted and actual energy consumption in buildings. By leveraging semantics, the system improves the accuracy of energy consumption predictions and provides stakeholders with valuable information. Integrating ANN and GA enables adaptive modelling, whereas rule-based semantic techniques improve interpretability. The proposed system can contribute to more efficient and sustainable building practices.

McGlinn et al. [237] presents a usability evaluation of an intelligent monitoring and control interface known as BuildVis, an integral component of a building energy management system (BEMS). This study aims to assess the usability of BuildVis in building energy consumption using an intelligent algorithm and a semantic knowledge base.

BuildVis incorporates an ontology based on Industry Foundation Classes (IFC) and uses Artificial Neural Networks (ANN), Genetic Algorithms (GA), and data mining techniques. The system employs these technologies to generate rules, select relevant rules, and provide energy-saving suggestions to Facility Managers (FM). Visualisation of the system is facilitated through a web-based interface developed using HTML5.

BuildVis' usability evaluation was conducted in five European buildings, targeting technical users and Facility Managers. The assessment focussed on gathering insights and lessons learnt regarding the system's usability in real-world scenarios. This study aimed to identify strengths, weaknesses, and areas for improvement to improve the user experience of the Web-based tool.

By evaluating the usability of BuildVis, this study contributes to holistic energy management of buildings. This study highlights the importance of incorporating intelligent algorithms and semantic knowledge bases to optimise building energy consumption.

2.8.3 BIM data vs BIM knowledge

It is essential to emphasise the distinction between building information modelling data and building information modelling knowledge. Building Information Modelling data has nearly doubled within the last decade, exhibiting an ever-increasing number of projects underpinned by BIM. Evolution is fuelled by exponential growth in computational power and hardware technologies, such as high-performance graphics cards that work in tandem to power BIM-based applications. Taking into account the crucial aspects involved in a BIM project that could affect the structural integrity of the building and consequently lead to significant health and safety concerns, theoretical and practical knowledge is needed as the standard method of teaching BIM and interacting with web applications becomes ineffective.

Even when interacting with BIM models on the Web in a rich graphical format, concepts may need to be more specific and clear. With the advent of Augmented Reality (AR) and MR, the content embedded in BIM web applications is matched with greater ease for physical assets. BIM ontologies are well-established ontologies of rigorously structured vocabularies that describe and interlink BIM terms in BIM models. A component of a building can be described using simple and general terms (i.e., "partition wall", "roof", "door"). More specific terms, such as parts of each component or material, significantly expand the vocabulary of the BIM-related ontology. The question remains: *How can we link all these BIM-related ontology terms?*. Developing digital content requires considerable effort and involves unique challenges. As co-design and co-creation were closed as marketing tools, users moved from organising in focus groups to enjoying the flexibility of designing products through participatory knowledge sharing, a key component of developing BIM models embedded in web applications.

Over many years, the IFC model and file format have been developed and evolved using the popular STEP/EXPRESS modelling language. A parallel scheme with XML, which is more relevant for the web environment but is based on a traditional, legacy, entity-related model, has been standardised. Furthermore, given streamlined formats such as the now-ubiquitous JavaScript Object Notation (JSON), XML-based document-based information management is increasingly viewed as cumbersome and verbose. JavaScript Object Notation (JSON) enables objects and concepts to be identified and explored using a standard Web browser using URLs. Data can be serialised in various RDF-compatible formats, reducing obstacles to its use by web-based developers. The ISES project, conducted by Wicaksono et al. [375], which used Ontology to address interoperability issues in a building energy management system, has recognised the benefits of managing operational constructions from this approach.

IfcOWL is characterised by extensive capabilities, including the linking of building and material data, manufacturing data, processes, and GIS data. With conversion to ifcOWL, the use of BIM data in web systems is improved, and more work is being done to further improve this [278].

Building Information Modelling (BIM is a process for creating and managing digital representations of the physical and functional characteristics of places Eastman et al. [112]. It provides a rich set of spatially organised multidisciplinary semantic data for a built environment Succar [344].

The distinction between BIM data and BIM knowledge lies in the depth and complexity of the information. BIM data refers to raw structured information captured and stored within a BIM system, whereas BIM knowledge is the interpretation and understanding derived from these data Succar [344].

However, the initial conversion of the EXPRESS constructs to semantic web constructs using ifcOWL results in a complex knowledge graph that can overwhelm developers Pauwels and Terkaj [279]. By simplifying traditional modelling patterns and restructuring the knowledge graph, the number of nodes and edges required to model an ontology can be reduced, improving usability and efficiency Beetz et al. [37].

Furthermore, aligning the ontological representations of smart cities and their subdomains with ifcOWL allows the usage of BIM data in a larger landscape, enhancing interoperability among sectors such as AEC (architecture, engineering and construction), facility management, and smart cities. This convergence results in a richer and more holistic user experience, allowing users to take advantage of the knowledge of BIM for more comprehensive and informed decision-making Biloria [49, 50].

2.8.4 BIM and Semantic social media analysis

The researcher could find numerous attempts to use the BIM industry field as an application for the mining of social networks, such as the one of [391], who attempted to mine the BIM design logs to discover connections between the characteristics of social networks and the success of designers' production. Furthermore, Kassem et al. [184] attempted to identify the critical competencies of the BIM expert positions, which were selected based on their quotes and a review of their overlap skills. However, the researcher could observe a gap in the current literature, represented by the need for more social media extraction, with the scope of improvement for the BIM industry, particularly concerning BIM roles and skills.

Barison and Santos [31] undertook extensive research on the BIM roles and skills advertised in job ad descriptions, while also performing a comparative analysis of market demands and efforts of universities and other institutions to integrate BIM into the curriculum. He also highlighted the importance of the correlation between BIM roles and skills for companies that adopt management competency models.

Wu and Issa [380] advises preparation to boost BIM curving while acknowledging that the qualifications of recent graduates are insufficient to meet the demand for industry jobs. Instead, they propose that BIM education should train graduates to be ready to the degree that organisations can influence their BIM skills according to their needs.

Succar and Sher [345] has analysed how educational and organisational institutions have started to adapt their delivery systems to meet changing market demands. This manuscript was one of the first to introduce taxonomies and conceptual models to clarify the mechanisms for filtering, classifying, and aggregating individual responsibilities into a competency database. It also discussed the benefits of the competency-based approach for industry and academia. According to the authors, individual BIM skills are the personal traits, technical skills, and professional skills required by an individual to conduct a BIM activity successfully or to produce a result linked to BIM. These skills, results, or operations could be measured against performance criteria and obtained through growth, training, and education or improved through development.

Meziane and Rezgui [242] and Antons et al. [18] emphasise the expertise of BIM for managing buildings and how a team leader with solid BIM skills can significantly impact project success and teamwork. The construction industry enjoys recruiting future employees with both BIM technology experience and broad analytical knowledge.

This subchapter provides a conceptual grounding in BIM, laying the foundation to examine its convergence with semantics and implications for user experience as we progress through the literature review.

2.9 Broader Societal Influences on User Experience

While the preceding sections focused on UX challenges within specific built environment domains, the following section provides a brief look at broader technological and societal concerns shaping user experiences.

The final subchapter of this review of the literature delves into the critical issues posed by misinformation and cybersecurity threats on user perceptions, trust, and engagement. In an era where fake news and cyberattacks proliferate, understanding the impact of these phenomena on UX becomes essential for designing resilient and trustworthy digital experiences. This subchapter also investigates the psychological and cognitive factors influencing users' susceptibility to fake news, explores strategies for enhancing UX resilience against cyberattacks, and sheds light on designers' and practitioners' ethical considerations and responsibilities in safeguarding user experiences among these challenges.

In the days of fake news and hacking, traversing the user interface may require significant effort. Customers expect to be able to rely on the information they obtain from various online sources. However, they want the ability to defend themselves against any possible attack. As a result, user experience designers must consider both of these considerations when creating designs [385].

The current state of affairs makes the threat of false news real and immediate. Fake news can significantly impact public opinion due to the ease with which false information can be created and disseminated and because people are more likely to believe something if it confirms their beliefs. As the researcher has seen in recent years, this can have potentially fatal consequences with the rise of populism and the dissemination of false information. Therefore, it is vital to educate consumers about the hazards of fake news and to take precautions to protect their safety [108].

A technique for providing a more trustworthy user experience is to make it easy for customers to identify evidence of a source's validity. This objective can be achieved by employing successful tactics, such as providing links to the source material or performing source checks before uploading content. Another way to improve the user experience is to make it easier for customers to report erroneous information or suspicious behaviour. Setting up a reporting system that is easily accessible and makes it easy to identify information suspected to contain errors is one method to achieve this objective [264][398].

In various ways, cyberattacks can have a disastrous impact on the user experience of a web application. A web application can become inaccessible or sluggish to load if, for instance, it is subject to a denial of service assault. Suppose that an attacker gains access to the servers hosting the application. In that case, they may be able to modify the application's source code or data, which could result in difficulties or unexpected behaviour. Under certain conditions, attackers can seize complete control of the targeted application [399].

To mount the most effective possible defence against cyberattacks, developers must ensure that the web application has adequate security safeguards, such as adopting strong authentication and authorisation procedures, appropriately configuring firewalls, and periodically patching security vulnerabilities. In addition, web developers should consider using security tools such as web application scanners to identify any potential flaws in their code [14][139][343].

Cyberattacks can substantially negatively impact the user experience by obtaining personal data that can lead to identity theft or by installing dangerous software that can interrupt normal operations. Malicious software can sometimes severely destroy the hardware of a computer system. A cyberattack can cause a decrease in the performance and availability of services and loss or corruption of data. Cyberattacks can occasionally result in fraudulent conduct or the theft of sensitive data. Unprotected websites are susceptible to hacking, which could result in data loss or theft. If a cyberattack is effective, it can result in service interruptions that enrage customers and force them to seek alternative providers [282].

Developing a safe and reliable information system that is immune to attacks is difficult for many businesses and organisations today. A single multi-role platform with the capacity to provide a high level of security and safety, the ability to monitor internal activity in real-time, and the capacity to respond swiftly to external invasions have become a must. Because of the potentially catastrophic consequences of a successful attack, it is clear that such a system is necessary. However, creating and implementing such a system is a labour intensive undertaking that requires substantial investment in time, money, and specialised knowledge. To ensure that the system remains effective despite constantly changing threats, it is also necessary to perform routine monitoring and testing [232][236].

Virtualised environments could provide a potential solution to this challenge, as they are a complete security solution designed specifically for virtualised infrastructures that offer high levels of protection against all types of threats, guest operating systems, and applications running on top of them. This configuration could be achieved using the advantages offered by virtualisation technologies, such as centralised management and rapid deployment, while providing comprehensive protection against known and emerging threats [138].

One significant advantage of virtual machine-based security is that it offers a high degree of flexibility in terms of deployment. For example, it can be implemented as a "white box" solution where all components are deployed within the same server or as a "black box" solution where the security system is located outside the protected network. This differentiation allows organisations to tailor their security solutions to their specific needs and requirements [55].

In addition, virtual machine-based security systems offer several other benefits over traditional physical security solutions, including a significantly reduced cost of deploying and maintaining a security system, as they eliminate the need for expensive physical hardware and reduce the time required to deploy and configure the system [216].

Virtualisation can also feature increased flexibility, since a virtualised security system can quickly adapt to rapidly changing threats and needs. For example, new security rules can be rapidly implemented without reconfiguring physical hardware or changing network infrastructure [93].

In summary, the proliferation of misinformation, cybercrime risks, and privacy concerns on today's Internet shape user experiences in complex ways across sectors. Manipulative fake content damages trust in legitimate sources, while cyberattacks threaten the adoption of networked technologies. However, digital literacy initiatives, ethical design principles, security measures, and regulatory protections can help mitigate these challenges. As the research gaps highlighted indicate, applying semantic techniques in a transparent and participatory manner can further bolster reliability, safeguard rights, and restore user confidence in built environment systems. Developing domain-specific solutions informed by ethical analysis and user-centred measures against misuse will be vital to ensuring these emerging technologies promote empowerment overexploitation. Ultimately, a comprehensive understanding of the sociotechnical ecosystem is imperative to engineer semantic systems that improve experiences responsibly. Understanding the impact of these phenomena on UX becomes essential to designing resilient and trustworthy digital experiences.

2.10 Gaps identified within the reviewed literature

This section includes definitions of research gaps created by original research questions. Smart cities are discussed as a system that offers an additional value proposition, but is currently being prevented from being realised due to a lack of shared semantic referential and robust knowledge exchange.

Semantic technologies enable IoT and smart city concepts to be fully realised. The semantic web and ontologies are explored in detail, and the need for integrative application layer interoperability research is argued to enable smart city solutions to enhance their capabilities for data exchange and higher-order business intelligence.

2.10.1 Addressing Gaps in Semantics for Enhanced User Experiences in Building Information Modelling (BIM) Systems

Building Information Modelling (BIM) is an essential component of the Architecture, Engineering and Construction (AEC) industry, serving as a crucial framework for data exchange. Despite its pivotal role, challenges persist in harnessing semantic technologies to augment user experiences in BIM-related training and learning environments. Recent studies have underscored the ongoing difficulties in knowledge transfer and the broader adoption of BIM methodologies. The potential of ontologies and formal knowledge representations to structure and contextualise information, thus facilitating the upskilling of personnel in diverse roles, still needs to be explored. This scenario opens avenues for developing customised semantic knowledge graphs, which could align competencies with tailored learning content and assessments, considering the individual's specific needs and background. More user studies are needed to evaluate semantically enhanced interfaces and content within BIM education, although preliminary prototypes indicate encouraging results. Future research should focus on developing intuitive visualisations and analytics strengthened by semantic inferencing to enhance understanding and reduce the barriers to engaging with BIM systems. Furthermore, adopting semantic standards for interoperable training content and learner profiles could significantly reinforce the ecosystem for persistent, personalised and shareable BIM knowledge.

2.10.2 Bridging Gaps in Semantic User Interfaces for Energy Management

The significance of Digital Twins in the built environment is increasingly recognised. Nevertheless, the intersection of semantic modelling, digital replicas, and adaptive user interfaces in energy management systems needs further research.

The potential of context-aware capabilities, enabled through semantic representations of building artefacts and real-time sensor data, is yet to be fully realised in creating bespoke user experiences. This thesis presents an opportunity for the application of participatory design methodologies, inviting energy system operators and occupants to collaboratively develop semantically rich interfaces that resonate with their specific needs and contexts.

A critical area for further exploration involves the integration of diverse data sources into ontology-driven Digital Twins, which could pave the way for responsive control interfaces and actionable visual analytics. Empirical research quantifying the impact of semantically transparent systems on user productivity, satisfaction, and adoption could strengthen the case for these emerging technologies. Although initial trials have shown potential benefits, considerable research and development efforts are essential for the wide-scale adoption of semantic-driven interfaces in energy management.

However, many businesses need more organisational intelligence to fully implement the steps that would allow them to benefit from the capabilities of such novel appliances. Deep Digital Twins, such as CUSP, seek to address this issue by providing an easy-to-use, intuitive user interface that allows stakeholders at all levels to leverage the power of semantics and ontologies in the ever-increasing management demands of a Digital Twin.

Rather than viewing a digital twin as a service for many industries and applications, the researcher focuses on the technology. Most studies on digital twins have focused on the manufacturing industry and related case studies to validate its design architecture. This previously identified research gap is being addressed by introducing a user experience-focused DTaaS (Digital Twin as a Service) that references semantic context, allowing users to benefit from a highly tailored approach. The researcher illustrates this approach as part of the work in 5.

2.10.3 Expanding Semantics in Urban-Scale Digital Twins

The field of urban informatics is increasingly interested in the use of city data to address the challenges of sustainability, mobility, and quality of life. Nonetheless, more research is required on methodologies and frameworks for developing city-wide Digital Twins that employ semantic techniques to integrate and give context to multisector data.

A promising prospect exists for amalgamating citizen inputs, gathered through social media mining and crowdsourcing, with official municipal data. Advanced semantic reasoning could provide insights on urban events, environmental conditions, and community sentiments. Employing predictive analytics on such a knowledge graph can provide urban planners with tools for participatory scenario simulations, enabling a systematic assessment of policy alternatives.

Despite the availability of geospatial ontologies, there is a need for greater standardisation to ensure interoperability across various urban systems. Usability studies can offer insights into how different stakeholders interact with urban digital twins, facilitating tailoring interfaces and visualisations to diverse user requirements. As urbanisation continues to escalate globally, more research on semantic urban twins is imperative to create more responsive and resilient cities.

Although the Internet of Things has grown in popularity in recent years, many businesses need more organisational intelligence to implement the necessary steps that would allow them to fully benefit from the capabilities of such novel appliances. Deep Digital Twins such as CUSP, underpinned by system-level control theory, seek to address this issue by providing an easy-to-use, intuitive user interface that allows stakeholders at all levels to leverage the power of semantics and ontologies in the ever-increasing management demands of a Digital Twin. This platform combines sensor and control data from digital models while keeping track of the interconnections between zones, various pieces of equipment, distribution systems, energy, and sensors associated with these components.

Therefore, there are additional levels of representation in addition to the classic real-world use case of a Digital Twin (modelling buildings and their associated characteristics, eventually leading to a business case derived from a reduction in non-renewable energy sources). Digital Twins could eventually represent their citizens' "alter ego" by transforming inhabitants into active agents in a smart city context (whether voluntary or not) and involving them in stakeholders' and policymakers' decision-making processes. This goal can be achieved using social media sources for text mining, event discovery, and sentiment analysis, as described in 6. This thesis aims to fill the intuitiveness gap by providing a much less invasive approach than mobile crowd-sensing applications through a mobile ecosystem.

Security was not addressed because the thesis focuses primarily on enhancing user experience rather than ensuring user safety. The researcher has created an iterative architecture for user experience, including security as one component.

The literature review has explored existing research on leveraging semantic techniques to enhance user experiences in applications for the built environment, focusing on building information modelling, energy management systems, and smart city digital twins. The review revealed several gaps and opportunities, including the need for further research on semantically enriched systems tailored to construction education, participatory design of ontology-driven interfaces for energy management, and integration of heterogeneous urban data sources into city-scale knowledge graphs.

Bridging these gaps requires adopting appropriate research methodologies. As the following chapter delineates, this research employs a mixed-method approach that combines qualitative techniques such as interviews and focus groups with quantitative data analysis and model development. A design science framework underpins the overall research methodology, allowing a rigorous process of artefact design, application in real-world contexts, evaluation, and refinement.

Case studies in domains such as BIM-based learning, energy management user interfaces, and urban digital twins provide real-world settings to apply and validate the proposed semantic solutions. Usability assessments provide vital data and user feedback to iterate prototypes.

2.11 The meaning of User Experience

Bridging the gaps illustrated in the previous section in leveraging semantic techniques to enhance user experiences requires clearly defining what constitutes a positive user experience in each application context and how it can be effectively measured. As the following sections elaborate, the user experience manifests itself in various ways across the domains of BIM-enabled construction training, energy management systems, and urban digital twin platforms. For each area, the user experience will be characterised by a targeted set of metrics and evaluation approaches tailored to the end users, their goals and the nature of interaction.

A combination of quantitative performance indicators and qualitative feedback will provide a comprehensive view. This framework to assess user experience in a contextualised, multifaceted manner will support the methodology to demonstrate and validate the benefits of applied semantic interventions in enhancing built environment systems.

This subsection illustrates the user experience definitions and considerations across the three core topics:

- BIM-based construction education
- energy management systems
- provides a solid framework to guide the UX research in this thesis

The multifaceted approach aligns with the overarching research objectives of enhancing user experiences through applied semantic techniques in built-environment systems.

The focus on usability, decision support, productivity, technology acceptance, and participatory engagement across user groups directly addresses the aim of improving utility, adoption, and outcomes. The emphasis on mixed evaluation methods incorporating subjective feedback and co-design also supports the goal of human-centred, experiential research.

Additionally, the UX facets identified help bridge critical research gaps highlighted in the literature review. Learner-centred design can aid in the acceptance of BIM training by the construction workforce. Semantic transparency in energy interfaces can improve operator situational awareness. The participatory features of digital twin enable citizen contributions to urban planning.

2.11.1 BIM for Construction Education

In the context of BIM-based learning platforms, user experience (UX) refers to the perceptions, attitudes, responses, and overall satisfaction of learners interacting with the system to acquire construction knowledge and skills. As learners engage with BIM training content and interfaces, their UX is shaped by multiple factors, including usability, engagement, pedagogical design, and technology acceptance.

Usability is a crucial determinant of UX, encompassing the ease, intuitiveness, and predictability of navigating and using various platform features. Metrics to evaluate usability include task completion rates, time on task, error rates, and subjective user satisfaction ratings. The usability of 3D model visualisations, navigation tools, and editing interfaces directly impacts learners' training efficiency.

The engagement of learners with platform content and peer collaboration spaces also shapes UX. Metrics for behavioural engagement can be quantified through the depth of content accessed, time spent, contributions made, and frequency of system use. Surveys and interviews gauging interest, focus, and perceived stimulation can provide self-reported cognitive and emotional engagement measures.

Pedagogical design factors, including perceived relevance, scaffolded challenge, and feedback, also contribute to training UX. Evaluations should assess training content's sequencing, scope and multi-modal delivery from a learner's perspective. Assessments can also determine knowledge gains, evaluating the platform's pedagogical effectiveness.

Technology acceptance factors like perceived usefulness, ease of use, and intention to adopt influence receptivity to BIM learning platforms, shaping overall UX. Standardised questionnaires can quantify acceptance and qualitative feedback on barriers and enablers to adoption. This multifaceted explanation provides a nuanced overview of critical elements characterising user experience in BIM training systems and how UX assessment should adopt a mixed methods approach combining multiple techniques tailored to each facet.

2.11.2 Energy Management Interfaces

For intelligent energy management platforms utilising semantic modelling and digital twin capabilities, user experience (UX) encompasses the situational awareness, cognitive load, decision-making, productivity, and overall satisfaction of operators and building managers interacting with the system.

Situational awareness is a critical facet of UX for energy control interfaces, focused on comprehending real-time building performance through data visualisations. Evaluating situational awareness involves measuring operators' ability to correctly interpret sensor data, identify anomalies, and determine current energy utilisation status. Both subjective questionnaires and objective comprehension assessments can be used.

Cognitive load represents the mental effort to assimilate and act on energy analytics dashboards.

In decision-making, UX involves clarity, transparency, and assistance in identifying and selecting appropriate energy interventions. User interviews, think-aloud protocols, and option tracing can evaluate the quality and process of decisions supported by the platform.

Productivity enhancement enabled by the energy management interfaces can be quantified through time savings in oversight workflows, reduced human effort through automation, and improved optimisation outcomes.

In general, user satisfaction surveys and focus groups that measure receptivity provide subjective, yet multidimensional, UX feedback to complement objective performance data. This combination of measurement approaches provides a comprehensive assessment of UX for semantically powered energy management platforms.

2.11.3 Urban Digital Twins involving citizens as active sensors

For city digital twins aimed at supporting urban planning and governance, user experience (UX) refers to the perceptions, acceptance, and usability feedback of diverse stakeholders interacting with the digital replica.

For policymakers and city officials, UX centres on the digital twin's actionable insights, situational awareness, and decision support. Valuable metrics include depth of exploration, recall accuracy when querying city indicators, and the perceived value of analytics for informing policies and resource allocation.

UX for urban planners is related to the intuitiveness and flexibility of tools for modelling and simulating development scenarios. Usability can be assessed by observing planners using features for traffic forecasting, infrastructure capacity planning, zoning regulations, etc. Feedback on workflow integration and comparison with conventional planning is also valuable.

Community members experience digital twins through public-facing visualisations and participatory features. UX metrics involve engagement levels, comprehension of projected impacts, and a sense of inclusion in planning. Surveys, social media sentiment, and focus groups can provide citizen perspectives.

Developers and technology partners focus on the extensibility, interoperability, and integration capabilities. Expert reviews of technical documentation, APIs, and ontology alignment tools can evaluate UX aspects such as learnability and transparency.

A comprehensive assessment of digital twin utility and usability across the urban ecosystem can be achieved by tailoring UX evaluation to each user persona's needs and priorities. A mixed methods approach provides both objective performance data and subjective user perceptions.

As cities deploy sensor networks and accumulate data, citizens can also serve as crucial "soft sensors" providing subjective lived experiences and ground truths to validate digital twin representations.

Citizen experience (CX) should be a key component of UX evaluations for urban digital twins. Surveys, interviews, and crowd-sourced reports can capture citizen perspectives on digital twin simulations' local relevance and accuracy. Feedback can identify missing dimensions of community life that have not yet been reflected.

For example, pedestrian activity mapping based on footfall sensors may not adequately reflect disabled access barriers or safety concerns. CX data could enrich this layer with mobility challenges and risk hotspots based on daily experiences of citizens.

Sentiment analysis and geotagged social data enable the incorporation of CX aspects like community belonging, trust in government, and neighbourhood vibrancy. CX metrics provide vital contextual and cultural insights that complement hard sensor data.

Participatory features allow citizens to validate and enhance the digital twin interactively through co-creation. Tracking contributions helps assess participation. The feedback of CX also shapes the iterative refinement of public interfaces to improve usability.

CX provides a critical subjective perspective on digital twin fidelity and indicates how successfully these tools promote civic participation. Incorporating citizen input creates richer digital placemaking reflecting lived urban experiences.

Through the methodology illustrated in the next chapter, the thesis aims to expand the scope and enhance the rigour of investigations at the intersection of semantics and user experience in the built environment.

3 Methodology

This chapter discusses the general methods used to complete the thesis. It will discuss the many stages leading up to the study presented in this thesis and the concepts and methodologies employed. A more extensive description of each research topic will be offered, along with the exact approach taken to address each topic and validate the suggested contribution. Toward the end of the section, the fundamental theory behind the primary strategies employed throughout this thesis is discussed. These techniques include semantic web modelling, three-dimensional web interface rendering, and semantic model sectioning.

To take a step towards filling the gaps identified in the literature review through a rigorous scientific process, this chapter describes the research design and methodology adopted. First, the philosophical and pragmatic paradigms under which the methodology is based are discussed. The research approach is based on established research design practices and modern developments to justify the processes undertaken to address the literature gap identified and framed through the research questions posed in Section 2.10.

This subsection starts by discussing the adopted high-level research approach, followed by the strategy and more detailed choices, before describing the pragmatic processes undertaken and the roles of these in answering the research questions. The researcher identified three distinct stages involved in the development of a smart city solution, corresponding to different stages of maturity:

- Theoretical Study
- Participatory action research project engagement
- Generalisation and unification based on Design Research

The role of each stage is described within the overall research design, along with their systematic methodologies. Their findings are then expanded in the following chapters.

3.1 Epistemological and Philosophical Perspective

The terminology "research philosophy" invokes a system of beliefs and assumptions about the development of knowledge. To provide stable and robust justifications for selecting the most suitable approach and strategy for research, it is fundamental to understand the underlying philosophy that dictates the worldview of researchers [323]. Every stage of research implies several methodological assumptions that every researcher makes [64].

There are different types of assumptions about the world and how we, as researchers, perceive the world through the lens of empirical evidence.

- epistemological assumptions, which primarily refer to human knowledge
- ontological assumptions, referring to the truths encountered within the research process
- axiological assumptions, referring to how the directions and values impact the research process.

These premises unavoidably shape how we understand research inquiries, use techniques, and decipher the findings [326]. All researchers depend on the reflexivity skills that allow them to become aware and actively shape the relationship between the philosophical position adopted and the way they carry out research [13]. Many researchers prefer the term "paradigm" rather than "research philosophy". However, both terminologies revolve around the same central idea. According to

Elgazzar and El-Gazzar [115], a ”paradigm” describes how a researcher shapes a worldview. He also defined the main paradigms of social science research, namely ”positivism” and “interpretivism”.

It has been argued that research philosophy and methods should not be viewed independently and that research philosophy should be addressed pragmatically, mixing models as required to answer the specific research questions posed [67]. It follows that enshrining a single philosophical doctrine regarding research design can harm the varied insights elicited from research practice. The views of Saunders et al. [321] on philosophical positions are broadly echoed by Ritchie and Lewis [303], who explain the role of philosophical paradigms as the basis of research design. Both sources identify positivism as the most clearly defined and acknowledged position, which holds that objective truth exists and should be gained through observation and empirical analysis, with the researcher remaining objective throughout. It contrasts with post-positivism, which acknowledges a tendency towards lack of neutrality of the researcher and attempts to account for this in research design. This further contrast is called antipositivism (also called “interpretivism” by Saunders et al. [322]), which says that objective truth should not be the research goal. However, reality and truth are experienced and interpreted by individuals [95].

Interpretivism combines schools of thought, particularly “constructivism” and “phenomenology” [95, 322]. As an epistemological philosophy, constructivism argues that the research process hinders the objectiveness of reality. Ritchie et al. [303] states that researchers can choose to remain neutral or participate personally in a study. Ritchie et al. [303] states that natural science research methods are inappropriate for deriving knowledge about the social world. This statement seems to apply within the realms of research, given the critical social dimensions of smart cities present at all stages of their lifecycle, whereby social and human-machine interactions are the central aspect of testing the hypothesis.

Despite the division of schools of thought, Saunders et al. [322] state that considering philosophical paradigms should not be viewed as a “shopping list” from which to choose the best option. Instead, they argue that understanding the implications of philosophical stances on research design is valuable in enhancing the production of valuable knowledge within a particular field. To this end, given the varied social and technological aspects of the research questions considered in the current study, the paradigm of pragmatism has significant value. It is again described by Saunders [323], who state that choosing and enshrining a paradigm may not be ideal and that the philosophical stance adopted should be tailored to the research questions. For this reason, the authors support the role of mixed-method research, engaging with quantitative and qualitative data. However, Orlikowski and Baroudi [274] argues that the addition of a traditional positivist perspective with interpretivism produces better research endeavours.

3.1.1 Summary table of approaches

To provide an overarching outline of the key research design choices made in this thesis, aligned with Saunders’ research onion model, the following table sums up the philosophical foundation, approaches, methods, and techniques adopted:

Research Onion Aspect	**Approach and Justification**
Philosophy	**Pragmatism** - Combining positivist and interpretivist stances provided the flexibility needed for multifaceted research questions.
Approach	**Mixed methods** - Both qualitative and quantitative techniques were required to gather user insights and evaluate prototypes.
Strategy	**Iterative prototype development** - An incremental build-test-refine approach allowed systematic enhancement of solutions.
Choices	**Mono method qualitative** - Interviews, focus groups, participatory design sessions. **Mono method quantitative** - Surveys, usage metrics, performance indicators. **Multi method** - Combining qualitative feedback with quantitative data analysis.
Time horizon	**Cross-sectional** - User evaluations made at specific points. **Longitudinal** - Tracking user interactions over time.
Techniques & procedures	**Literature analysis** - Identifying research gaps. **Requirements gathering** - Understanding user needs. **Prototype development** - Building semantically enriched solutions. **Usability testing** - Evaluating prototypes through usage. **Surveys & interviews** - Getting direct user feedback.

Table 1: Summary of research philosophy, approaches, and methods

This table presents a concise yet comprehensive overview of the methodology, encapsulating the pragmatic philosophy, the mixed-method approach, the iterative strategy, specific procedures, and techniques used in the research. It provides a high-level contextualisation of the detailed methodology delineated in the subsequent sections.

As illustrated in the summary, this thesis embraces a pragmatic paradigm that guided the use of various qualitative and quantitative techniques that are suitable for research questions. Iterative prototype development enabled incremental enhancements based on user feedback. Cross-sectional and longitudinal aspects provided complementary usage insights over time.

This foundation established, and the following sections delve deeper into the specifics of the methodology, connecting the philosophical perspectives to the practical research design and process. The mapping to Saunders' layered model situates the research techniques within validated frameworks for business studies. By outlining the methodology in a structured, yet flexible manner, the research aims at rigour and relevance in addressing the challenges of applying semantic techniques to enhance user experiences in the built environment.

3.2 Research Design

3.2.1 Research approach and strategy

The research design embraces an interpretative approach that aligns perfectly with the philosophical stance of the researcher and the study objectives. Interpretative approaches are widely recognised in social science research, as they offer a platform for understanding social constructs and phenomena, focusing on meaning and human experience [230].

The use of an interpretative approach in the outlined research design is well suited and pivotal for several reasons. Firstly, the research is not only focused on the technicalities of the BIM web applications but is also deeply interested in understanding the nuanced relationship between semantics and the user experience. This requires a research approach that delves beyond these applications' surface-level functionality to understand the user's subjective experiences and interpretations.

The interpretative approach in this research enables the examination of how users interact with and make sense of enhanced BIM web applications, how they engage with the incorporated semantics, and how these semantics consequently influence their learning experience, energy management practices, and their role as active sensors in the smart city context. Such insights cannot be gleaned through purely positivist or empirical approaches; the interpretative approach provides a deeper and more nuanced understanding of the subjective realities of the users.

An interpretative approach is well aligned with the iterative and reflective nature of the research design. Each stage of the research builds on the understanding of the previous stage, which requires a research approach that acknowledges the evolving, context-specific, and interpretative nature of human behaviour and experiences [244].

The researcher, rooted in an interpretative philosophical position, has designed a robust and comprehensive methodology guided by the general steps of literature review, gap analysis, prototype implementation, and continuous feedback integration. This methodology ensures that at each stage of the research, the researcher is not just observing the users' interactions with the BIM applications but is also engaging in an active process of interpretation to understand the underlying meanings and contexts. These interpretations guide the refinement of prototypes and the design of subsequent stages, embodying the iterative and reflective spirit of the interpretative approach.

In the context of this research design, the interpretative approach offers a nuanced understanding of how semantics enhance the user experience of BIM web applications. This interpretative lens enables the researcher to unpack the complex interplay between semantics, user experience, and BIM web applications, making this approach suitable and integral to this research.

Building on the philosophical perspective and research approach adopted by pragmatism-orientated multimethod research, this section describes the experimental processes and actions undertaken and their time frames to answer the research questions.

Figure 6: Methodology Research Design Steps

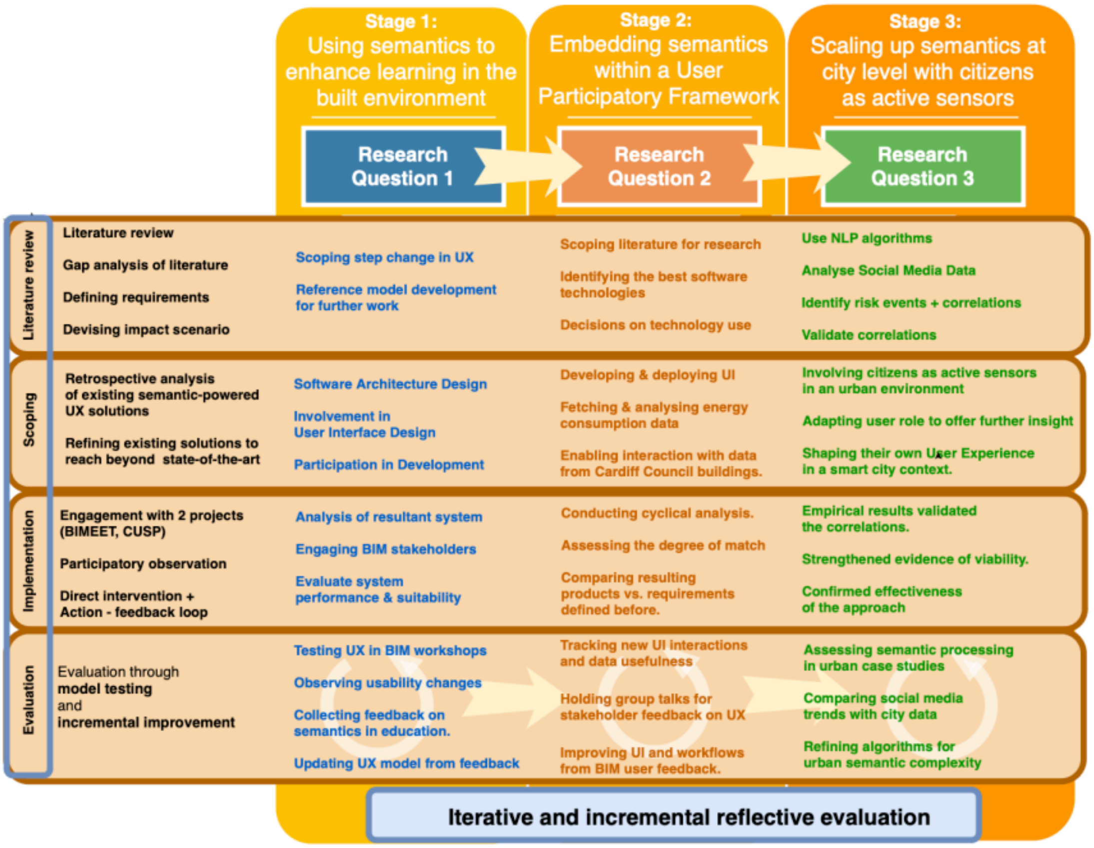


Figure 6 illustrates the research framework underpinning this thesis. It proposes an iterative and incremental reflective approach, which involves (a) using semantics to enhance learning in the built environment, (b) embedding semantics within a User Participatory Framework, and (c) scaling up semantics at the city level with citizens as active sensors. The three stages are illustrated within the vertical ellipsis, along with the three research design interventions revolving around literature reviews, gap analysis, implementation of prototypes or enhancing existing products, and seeking and implementing feedback iteratively.

Each stage forms a key component of the research design, intersecting with literature reviews, gap analysis, the development and enhancement of prototypes, and the continuous solicitation and integration of feedback.

1. Stage I (Using semantics to enhance learning in the built environment) specifically addresses the first research question: *How could an existing web application be enhanced using context in a way that maximises User Experience?*

2. Stage II (Exploring the role of semantics in enhancing UX through a Web-based User Interface for an Energy Management Platform) addresses the second research question: *How to derive context from static artefacts (buildings) and dynamic artefacts using a Digital Twin?*
3. Stage III (Scaling up semantics at the city level with citizens as active sensors) addresses the third research question: *How can we rely on this diverse (i.e., tangible and intangible) dynamic context to enhance a user's experience (e.g., a stadium, city planner, city manager) in a wide range of applications applied in the built environment?*

The research approach delineated in this thesis represents a meticulous blend of inductive and deductive methodologies, meticulously tailored to address the nuanced research questions and objectives identified at each stage. This hybrid methodology ensured a comprehensive and rigorous exploration of the complex domains within which the research was located.

The approach was primarily inductive for the initial research question, which focused on improving an existing web application using contextual data. The exploratory phase was based on a detailed analysis of current industry practices and user feedback within construction training platforms. The objective was to uncover latent opportunities for semantic enhancement, which were then actualised through prototype development and iterative refinement. This process epitomised the inductive method: observations and specific data points guided the gradual formation of broader theories and solutions.

On the contrary, the second research question, which probed the derivation of the context from static and dynamic artefacts, necessitated a more deductive approach. Drawing upon established theories of digital twins and semantic modelling, the research ventured to conceptualise and construct the CUSP platform with a clear, goal-orientated agenda. The research was based on the hypothesis that the integration of semantic context from digital twins could significantly enhance the user experience. Thus, the deductive method was employed, applying theoretical constructs to precision engineer the energy management prototype.

The investigation into the third research question, concerning applying semantic techniques at the urban scale, reverted to an inductive stance. Data collated from citizen sensors was the foundation for detecting emergent patterns and deriving insightful theories. The inductive strategy was evident as exploratory data gathering and multiple regression analysis revealed statistically significant correlations, fostering theories related to urban events, environmental factors, and citizen satisfaction. Regarding overarching objectives, the development of semantic web mining for construction training embarked on an inductive journey, sifting through data to delineate roles and requisite skills. However, the creation of the CUSP platform was emblematic of a deductive trajectory, applying established semantic and digital twin theories to a concrete end. The endeavour to refine urban planning via participatory modelling adhered to an inductive pathway, with citizen inputs instrumental in shaping the development of pragmatic tools.

The thesis's methodological fabric is interwoven with both inductive and deductive threads, each selected and applied precisely based on the research stages' specific demands and aspirations. This adaptive and contextual synthesis of inductive and deductive reasoning fortified the methodological integrity needed to navigate and fulfil the multifaceted goals of the research.

3.2.2 Developing a new User Experience model

Developing a new model for user experience (UX) in the realm of semantic web applications within the built environment sector is both a necessary and insightful academic pursuit. This is primarily because existing UX models do not adequately cater to the specific requirements of web development, particularly concerning the unique complexities inherent in the semantic web and the built environment. The proposed model is a response to this gap, with the aim of providing a comprehensive framework that aligns with the distinct characteristics of this emerging field.

Central to the necessity of this new model is the recognition that traditional UX theories and frameworks, while foundational, fall short in addressing the intricacies of semantic web applications in contexts such as construction, architecture, urban planning, and the development of smart cities. The seminal works of Hassenzahl, Mahlke, and Karapanos are invaluable; however, they are only partially applicable to the nuanced requirements of this specialised domain. Consequently, the proposed model seeks to extend and adapt these established theories to better fit the semantic web's unique characteristics within the built environment.

Three primary areas of focus underpin this new model:

- **Technical complexity:** Semantic web technologies introduce a level of technical sophistication that is not typically found in conventional software applications. This complexity is characterised by advanced knowledge representation and reasoning. The proposed model recognises the need to account for the cognitive load and learning curves associated with these technologies, as well as the interoperability challenges that arise in knowledge-driven systems.

- **Domain knowledge:** In the built environment, domain-specific knowledge is imperative, especially when dealing with complex data related to architecture, engineering, sustainability and urban planning. The model emphasises the importance of considering users' domain expertise and how it influences their motivations, expectations, and ability to derive value from semantically enriched systems.

- **Stakeholder diversity:** Semantic Web applications in the built environment must cater to a wide range of stakeholders, including architects, city planners, building owners, facility managers, and citizens. The model underscores the importance of designing inclusively for these diverse user groups, ensuring that the needs of all stakeholders are met.

- **Technical complexity:** The proposed model integrates additional constructs derived from existing theories, such as the Technology Acceptance Model (TAM), Cognitive Load Theory, and User-Centred Design (UCD) principles, to address these challenges. The model assesses technical acceptance through perceived usefulness and ease of use, assesses domain knowledge to identify learning needs, and employs stakeholder empathy to fully understand the diverse user requirements.

The model presents a robust and context-specific framework for evaluating and enhancing UX in semantic web applications tailored to the built environment by weaving these elements with foundational UX principles like usability, motivation, emotion, and value. This fusion of established UX theories with new considerations specific to this technology domain creates an academically solid and practically relevant evaluation tool.

Moreover, the model's criteria are interdependent, enhancing its robustness. For example, technical acceptance is influenced by users' prior domain knowledge and the empathetic design tailored to stakeholder needs. This multifaceted approach allows for a more holistic understanding of UX, facilitating the development of more human-centric semantic web applications in the built environment.

The proposed model not only fills an existing gap in UX research, but also lays a foundation for future academic and practical advancements in this dynamic and evolving field.

This new UX model will serve as an underlying framework to guide learning enhancement in the built environment using semantics in Stage 1. Additionally, it will inform the participatory design process for the energy management platform in stage 2. Furthermore, the model will provide the basis for developing dashboards, visualisations, and participatory features for urban planners and citizens in stage 3.

A vital aspect of the model is that it will undergo iterative refinement based on user studies and participatory design input collected during each stage. Feedback from learners, building owners, urban planners, and other users will reveal additional UX factors unique to the specialised context of semantic web applications. These insights will be incorporated to evolve the model, ensuring that it continuously adapts to users' needs and remains contextually relevant.

The model will facilitate evaluating and optimising UX across all three stages by considering the specialised considerations for semantic web applications. The criteria of technical acceptance, domain knowledge, and stakeholder diversity will guide the development and assessment of the prototypes and platforms at each stage.

This new UX model will serve as an underlying framework to guide learning enhancement in the built environment using semantics in Stage 1. Additionally, it will inform the participatory design process for the energy management platform in stage 2. Furthermore, the model will provide the basis for developing dashboards, visualisations, and participatory features for urban planners and citizens in stage 3.

A vital aspect of the model is that it will undergo iterative refinement based on user studies and participatory design input collected during each stage. Feedback from learners, building owners, urban planners, and other users will reveal additional UX factors unique to the specialised context of semantic web applications. These insights will be incorporated to evolve the model, ensuring that it continuously adapts to users' needs and remains contextually relevant.

The model will facilitate evaluating and optimising UX across all three stages by considering the specialised considerations for semantic web applications. The criteria of technical acceptance, domain knowledge, and stakeholder diversity will guide the development and assessment of the prototypes and platforms at each stage.

As such, developing a customised UX model provides crucial theoretical grounding before embarking on the applied research stages and will serve as a unifying framework for elevating UX throughout the methodological process while evolving through user feedback.

In the journey towards crafting user experiences (UX) that resonate deeply with end-users, it is essential to synthesise established psychological frameworks with pragmatic product evaluation metrics. This thesis aims to forge a robust UX model for building information modelling (BIM) and city information modelling (CIM) web applications. Weaving the theoretical underpinnings of Hassenzahl, Mahlke, and Karapanos with the empirical grounding of the Technology Acceptance Model (TAM), this study constructs a comprehensive approach to assess and refine UX within the context of the built environment. This integration acknowledges the inherent complexities, user diversity, and evolving nature of interactions that shape user experiences over time.

In Figure 7 we can observe the components of our bespoke UX model and how they interact with each other.

Hassenzahl's model core components:

The foundation of the new UX model is built on the three core elements proposed by Hassenzahl: the *user*, the *system*, and the *context*.

The user component encompasses the internal state of the individual interacting with the system, including their predispositions, expectations, needs, motivations, and mood (also present within Hassenzahl's framework). The user profile shapes the lens through which they perceive and respond to the system.

In BIM and CIM web applications, users are diverse, encompassing architects, engineers, city planners, and the general public. The model needs to prioritise understanding their needs, behaviours, and how they interact with the technology.

The system component refers to the product, device, or service with which the user interacts. It includes the complexity, goals, usability, and functionality of the system, values which align with Hassenzahl's model. The characteristics and design of the system have a significant impact on the user experience.

BIM and CIM web applications encompass technical aspects such as interface design, functionality, and interaction mechanisms. By integrating this aspect, the model can assess how these aspects facilitate or hinder the user experience.

The context comprises the physical, social, cultural, and environmental factors surrounding the interaction. As per Hassenzahl's model, the context can transform the user experience even when the user and the system remain constant.

Within the context of BIM and CIM web applications, this refers to the specific environment in which they are used, including the built environment, urban planning scenarios, and various stakeholders' roles.

By integrating the context, which in the use cases of this research, could include the physical environment, the socio-cultural background of users, and the specific scenarios in which the applications are used.

Mahlke's Temporal/Dynamic Aspects:

To incorporate the temporal dynamics emphasised by Mahlke, the model proposes measuring the user, system, and context components at multiple points throughout the experience, capturing the evolution of the interaction and how the components change.

This aspect considers the evolution of UX over time, highlighting the ongoing interaction between user, system, and context. In the realm of BIM and CIM, this includes the changing needs and behaviours of users and technological advancements.

Measurement of UX over time allows us to assess learning curves, adaptation processes, and

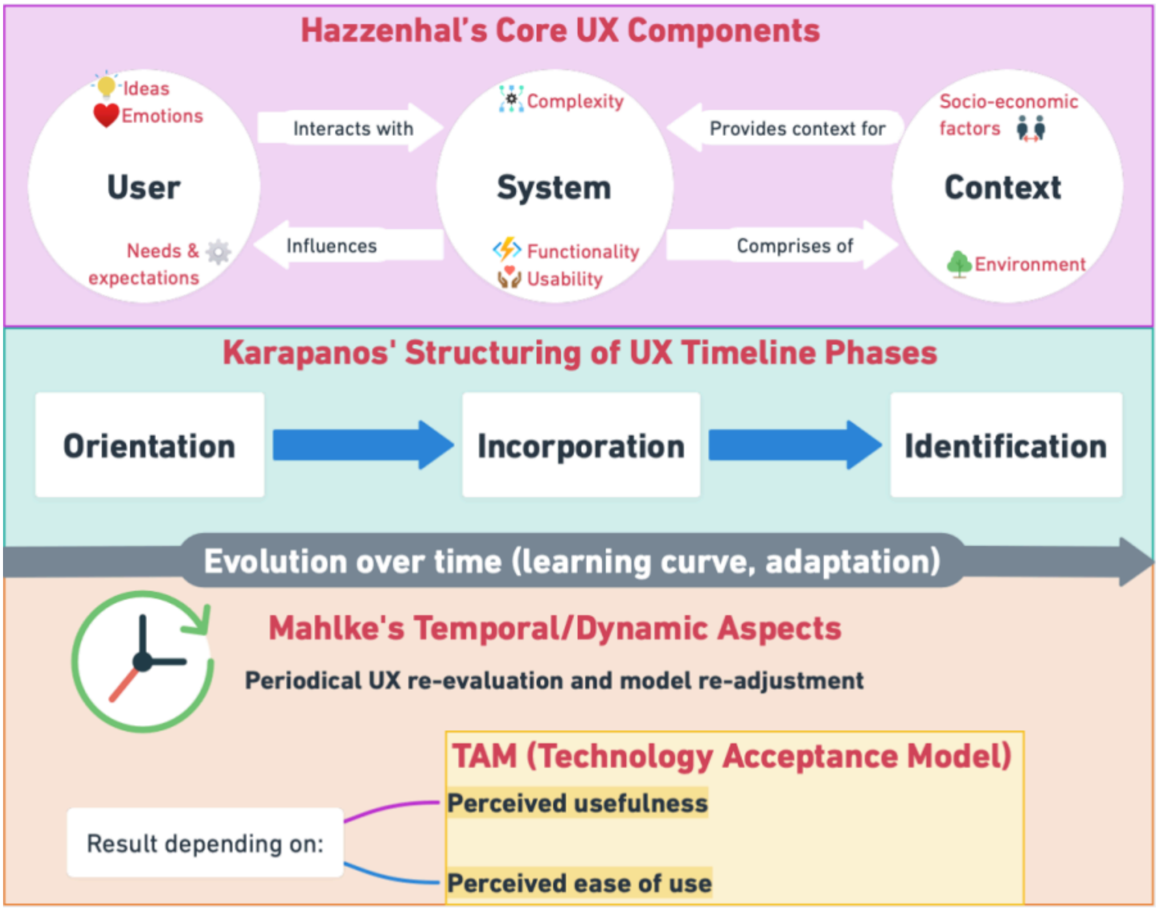


Figure 7: The bespoke User Experience Model

evolving user satisfaction.

Within the context of BIM and CIM web applications, Mahlke's emphasis on the UX interaction over time is integrated to observe how users' relationships with BIM and CIM applications evolve. This dynamic perspective allows the model to capture changes in user perceptions, preferences, and their adaptation to the system.

Karapanos' Structuring of the UX Timeline:

During the orientation phase, the focus is on novelty, with an ambivalence between excitement and frustration, as users initially become familiar with BIM and CIM web applications. The focus here is on initial impressions, learnability, and essential utility. This initial phase is crucial for first impressions and early learning. The model evaluates how users initially perceive and begin using BIM and CIM applications, focusing on aspects such as intuitiveness and initial barriers.

Users start integrating these tools into their regular workflows during the incorporation phase, where long-term usability becomes prominent as the product becomes a habit. This phase evaluates the deeper engagement, efficiency, and impact on productivity. As users become more familiar with the applications, this phase assesses how the tools are integrated into regular workflows, looking at efficiency, effectiveness, and satisfaction. In the long-term relationship with the system, this phase evaluates deeper emotional connections, loyalty, and how well the system aligns with users' evolving needs and goals.

During the identification phase, users form a long-term relationship with the system, developing emotional attachment and personal identification with the product. Loyalty is adopted if the system aligns with their evolving needs and goals.

TAM:

In addition to the core elements of user, system, context and temporality, the new UX model integrates metrics from the Technology Acceptance Model (TAM) to evaluate users' motivations to accept and use a technology based on two key determinants - perceived usefulness and perceived ease of use.

Perceived usefulness refers to the degree to which a user believes that using the system will improve their performance. Perceived ease of use refers to how effortless the user perceives the system. According to TAM, both these factors shape the user's attitude towards using the system, which, in turn, affects their intention to use it and actual usage behaviour.

In conjunction with Hazzenhal's model, TAM presents significant potential to enhance the model's capacity to understand and predict user interactions with BIM and CIM applications, ensuring a comprehensive approach to evaluating user experience in the built environment.

The emphasis of TAM's perceived ease of use and perceived usefulness aligns directly with the user's interaction with the system. This integration offers vital insights into how users perceive the utility and accessibility of building information modelling (BIM) and city information modelling (CIM) applications. It probes into the user's subjective assessment of the system, investigating how their perceptions of usefulness and ease of use influence their willingness to engage with BIM and

CIM tools.

On the system front, TAM's constructs are employed to evaluate how the design and functionalities of BIM and CIM applications influence user acceptance. This assessment is crucial because it determines whether the system's interface and capabilities align with user expectations and ease-of-use requirements. The integration of TAM here facilitates a nuanced analysis of the system's design, allowing for a deeper understanding of how various aspects like layout, navigation, and features contribute to or detract from the overall user experience.

In considering the context, TAM provides a framework to analyse contextual factors that influence technology acceptance, including aspects such as organisational culture and the specific requirements inherent in the built environment. By incorporating TAM, the model gains a richer perspective on how external factors, such as the workplace environment or sector-specific demands, play a pivotal role in shaping the user acceptance of BIM and CIM applications. This holistic approach ensures that the model is user-centric and responsive to the broader context in which these technologies are deployed.

In conjunction with Mahlke's temporal and dynamic aspects, it provides a longitudinal perspective in the research, understanding how user perceptions of BIM and CIM applications evolve. This integration allows tracking changes in user attitudes towards these technologies as they become increasingly familiar. By adopting this approach, the model captures the dynamic nature of user experience, offering insights into how initial acceptance of the technology, based on perceived ease of use and usefulness, may change as users become experienced using the system and their needs and expectations evolve. This longitudinal view is vital to comprehending the entire user interaction lifecycle with BIM and CIM technologies, ensuring that the model remains relevant and responsive to the changing nature of technology and user engagement over time.

In conjunction with Karapano's model, perceived usefulness and ease of use metrics are captured across the orientation, incorporation, and identification phases preserved from Karapano's model. Comparing these metrics temporally and with user, system, and context data enables a deeper understanding of what drives acceptance. For example, software glitches that affect the ease of use during orientation can hinder the incorporation of the system into daily habits. Usefulness perceptions may override early frustrations, facilitating long-term adoption. TAM data enriched by temporal and contextual information offers a powerful lens into user motivations.

During this initial orientation phase, TAM plays a crucial role in understanding users' first impressions and initial acceptance of BIM and CIM applications. It focuses on how users perceive usefulness and ease of use at the beginning, which are critical factors that influence their early engagement with the technology.

As users become more familiar with the system, TAM offers valuable insight into the evolution of their acceptance. This phase examines the factors that contribute to the increase or decrease in the usage of BIM and CIM applications. TAM helps identify the aspects of the system that either facilitate or hinder its integration into users' daily tasks and routines.

During the identification phase, TAM aids in assessing long-term acceptance and the sustainability of the system's use. It evaluates whether the BIM and CIM applications continue to meet user needs regarding usefulness and ease of use as users develop a more profound and long-standing relationship with the technology.

Through these phases, TAM provides a structured approach to assess user interaction with BIM and CIM applications over time, providing a comprehensive understanding of how acceptance, usage, and user satisfaction evolve throughout the lifecycle of technology engagement.

Considering all the aspects that TAM complement, it leverages its potential to make our model capable of analysing statistical relationships between TAM metrics like perceived usefulness and contextual factors like social influence. In addition, it examines correlations between perceived ease of use and system components such as complexity, all of which are valuable insights into acceptance and technology integration.

Integration:

In constructing this model, it is essential to integrate these elements seamlessly. The model begins with a clear understanding of the diverse user base and their specific contexts within the built environment. The system's design must align with these contexts while being adaptable to change over time.

The temporal aspects of Mahlke's model remind us that UX is not static; it evolves as users' familiarity with the system grows. This evolution must be captured through continuous feedback and iterative design processes.

Finally, Karapanos' phases guide us in structuring the UX assessment, ensuring that the model captures the full spectrum of user engagement - from initial acquaintance to long-term commitment.

In this research context, this model presents the following strengths:

1. It offers a comprehensive framework to evaluate and improve UX in BIM and CIM web applications. It aligns with the thesis's focus on leveraging semantics to improve UX, acknowledging the dynamic nature of user interactions and the multifaceted nature of the built environment.
2. It contributes to the academic understanding of UX in these specific domains and provides practical insights for practitioners in the field.
3. It is tailored to assess the unique characteristics and complexities of web applications in the built environment, recognising the importance of user-centred design, technical sophistication, and contextual relevance.

The model provides a comprehensive, multidimensional perspective on UX by assessing the user, system, and context components within each temporal phase. Questionnaires, interviews, and data logging can capture metrics for analysis.

Statistical analysis examines the correlations between the components and temporal variations. The results inform the iterative refinement of systems to improve UX. This new model holistically evaluates UX by incorporating Hassenzahl's elements, Mahlke's temporality, and Karapanos' phases. Multidimensional analysis supports tailored and user-centric design. The additional integration of TAM into this bespoke UX model offers a comprehensive framework for assessing user acceptance of BIM and CIM web applications, providing a nuanced understanding of how users perceive, use, and continue to engage with these technologies over time, ensuring the model's relevance and applicability in real-world scenarios.

A model grounded in well-established UX theories and adapted to the specifics of BIM and CIM solidifies its contribution to the body of knowledge in user experience, particularly in the context of smart cities and the evolving digital landscape of the built environment.

This re-imagined UX model is a confluence of psychological insights and methodological rigour crafted to gauge and enhance user interactions with BIM and CIM applications. This thesis solidifies a landmark advancement in UX research, charting a path for future exploration in smart cities and the dynamic digital landscape of the built environment. It extends beyond mere theoretical musings to provide a framework for real-world applications to be meaningfully evaluated, iterated, and refined.

A vital aspect of the bespoke UX model is its continuous refinement throughout the research process based on user studies and participatory design input. As the subsequent stages unfold, user feedback reveals additional factors unique to semantic web applications in the built environment context. These insights are incorporated to evolve the model, ensuring that it adapts to emerging needs.

3.2.3 Stage 1: Enhancing Learning in the Built Environment with Semantics

The potential of Industry Foundation Classes (IFCs) was examined to improve the learning experience of construction industry professionals [314]. IFCs, as an open and standardised data model, were identified as key to improving knowledge acquisition and comprehension in the built environment. This conceptual foundation laid the foundation for subsequent empirical exploration and evaluation.

Then a gap analysis was performed to assess the state of current learning mechanisms in the construction industry and identify areas requiring semantic enhancement. Following this, the researcher designed and implemented prototypes of enhanced learning platforms, infusing them with the identified semantics to enhance the learning experience. The prototypes were evaluated iteratively, incorporating user feedback and refinement at each iteration.

3.2.4 Stage 2: Embedding Semantics within a User Participatory Framework

The second stage aimed to refine processes by enhancing the user experience of a platform that manages and optimises energy consumption within a building. Based on the literature on semantics in user experience (UX) design and based on the semantics identified and tested in Stage 1, the researcher focused on improving the user interface and user engagement of the energy management platform.

To achieve this, the identified semantics were embedded in a User Participatory Framework [306]. This framework informed the enhancement of an existing energy management platform, incorporating user feedback and iteratively modifying the platform accordingly. This participatory approach facilitated the harmonisation of technological features with user preferences and needs, ensuring an enhanced and personalised user experience.

3.2.5 Stage 3: Scaling up Semantics at the City Level

In the final stage, recognising the limitations of focusing solely on building-level energy management, the researcher expanded the scope to include a wider city context, intending to integrate users (citizens) as active sensors in a smart city context [191].

After reviewing the current literature on semantics and smart cities, a gap analysis was performed to uncover unmet needs and potential improvements. Informed by these insights, the researcher designed and implemented a city-wide platform, using citizens as active sensors to provide location-specific real-time data on various city parameters. Citizen feedback was continuously sought and integrated into the platform, fostering a sense of community ownership while ensuring the continuous evolution and enhancement of the platform.

The research design approach sought to harness the power of semantics to improve the UX of Building Information Modelling (BIM) web applications through a three-stage, iterative, and reflective process. This methodology directly addressed each research question, and the methodological approach at each stage was informed and built on the results of the preceding stage. Through the integration of literature reviews, gap analyses, prototype development, and continuous feedback integration, this research design facilitated a rigorous and comprehensive exploration of the role of semantics in improving BIM web applications.

3.2.6 Transferability

The methodology devised within this research has been carefully designed to ensure adaptability and transferability across various domains, extending far beyond the specific case studies of building information modelling, energy management, and smart cities to which it was initially applied. Its fundamental construct is rooted in a versatile and robust framework that can be tailored to suit various sectors and problem contexts.

At the heart of this methodology lies an iterative approach encompassing comprehensive literature analysis, precise gap identification, innovative prototype development, and the integration of user feedback. This process, punctuated by its versatility, stands as a universal blueprint to enhance user experiences in many applications. The use of participatory design techniques within this research offers a tried and tested model for effectively weaving end-user perspectives into the fabric of system design, ensuring outcomes that are not only technologically sound, but also resonate deeply with user needs and expectations.

Moreover, the use of semantic modelling in this research is a powerful tool to structure complex data and extract meaningful contextual insights, showcasing broad relevance in industries grappling with intricate information systems. Coupled with the ability to interpret unstructured data sources through advanced natural language processing techniques, the methodology opens new avenues for harnessing rich open-ended user feedback, offering a valuable resource for continuous improvement and innovation.

As explored in this research, the concept of the digital twin holds immense potential for a wide array of applications where virtual representations of physical assets can significantly enhance monitoring, simulation, and decision-making capabilities. The urban-scale digital twin prototype, in particular, demonstrates the practicality of amalgamating multisectoral data sources into a cohesive and comprehensive platform, indicative of the methodology's potential to inform and transform practices in any context where digital and physical convergence is vital.

Although specific alterations would be necessary to refine and customise the methodology for new problem spaces, the core principles underpinning this research – including participatory and user-centric design, semantic modelling, digital twin integration, and iterative refinement, stand as pillars of a universally applicable framework. The prospect of conducting additional case studies in diverse fields such as healthcare, transportation, or manufacturing presents an exciting opportunity to further substantiate the transferability of the methodology. Through comparative implementations, we can uncover required context-specific adaptations, thereby enhancing our understanding of the methodology's versatility while preserving its fundamental essence.

3.3 Stage I - Using semantics to enhance learning in the built environment

At the beginning of the methodological journey, the primary objective aligns with the response to the initial research question: *How could an existing web application be enhanced using context in a way that maximises User Experience?* This question anchors the exploration within web application enhancement, focusing on the integration of context and its implications for User Experience (UX). This stage signifies a critical juncture in the investigation, establishing the foundation for subsequent stages and establishing the first pillar of the interpretative framework.

Context serves as the underpinning construct of semantics, since it encompasses the circumstances or facts surrounding a particular event, situation, or user interaction. In web applications, context plays an essential role in shaping the user experience, influencing how they interpret and interact with the application interface [82]. Therefore, by enhancing a web application using context, the researcher seeks to harness the potential of semantics to enrich the User Experience.

At this stage, the researcher delves into the intricacies of incorporating context into an existing web application. We strive to uncover the potential paths for maximising User Experience, and in doing so, the researcher employs a nuanced and meticulous approach. The process involves meticulous literature analysis, critical gap analysis of the current state of the art, and practical implementation of contextually enhanced prototypes. Iterative feedback, reflective of the interpretative stance of the research, is sought and incorporated, ensuring that the evolved prototypes resonate with users' expectations and needs.

By addressing the first research question, this research stage does not merely serve as a building block for the subsequent stages but also stands as a testament to the overarching goal of the research: enhancing User Experience through the power of semantics. Thus, as we navigate through this initial stage, we remain steadfast in the commitment to extracting meaningful insights and developing innovative solutions to maximise User Experience, setting the tone for the rest of the journey.

This stage of the research leverages text and social media mining, including material from texts and web content, to comprehend BIM's dependence on energy-related ideas mainly linked to roles and skills while presenting a methodology for classifying their frequency and correlations. The analyses are performed on a semantic web platform (*energy-bim.com*), supporting TF-IDF techniques and importance evaluation methods to identify BIM skills and roles with the corresponding dependencies. Additionally, at this stage, the researcher aims to understand the implications of the construction industry's digital transformation, including the gradual adoption of BIM, regarding required roles and skills to harmonise and devise training and educational programmes for the next generation of construction professionals.

The researcher has identified a clear literature gap, characterised by the need for studies that combine the analysis of secondary sources of evidence with information from social networks. The researcher used TF-IDF, association, correlation and cluster analysis to infer roles and responsibilities for adopting BIM in the industry.

The methodological stage uses a positivist philosophical stance using a mixed method consisting of quantitative and qualitative approaches to analyse secondary sources of evidence drawn from the academic and industry literature and social media sources associated with authoritative institutions and experienced practitioners. Based on consultation studies carried out as part of the EU H2020 BIMEET project [286], the researcher has adapted and exposed a semantic web platform, *energy-bim.com*, which can manage, store and analyse BIM-related information. The platform supports BIM knowledge sharing and enrichment of BIM knowledge within a community of BIM professionals and resources with the intention of advancing the implementation of BIM for energy efficiency in the construction sector. The *energy-bim.com* platform integrates various BIM-related data sources, based on which a set of TF-IDF and Metric Cluster methods is applied to determine relevant roles and skills around BIM.

3.3.1 BIMEET - Platform overview and motivations

BIMEET is an example of scaling up the BIM field to include other aspects through the addition of new concepts. It forms the basis of the second generation of web portals that underpin semantics that form part of it.

BIMEET endeavours to enhance the skills, qualifications, and capabilities of construction practitioners (from high-level professionals to blue collar workers), thus increasing market penetration and adoption of key technological development in BIM, given the timeliness of the need for training in combined green and functional performance engineering. Several areas are key to the potential growth of BIM for energy efficiency and its impact on the green building marketplace.

1. The multidisciplinary integrative capacity of BIM: BIM provides a unique opportunity to integrate data, information and underpinning processes across life-cycle and supply chains. This will promote informed and energy-efficient design interventions.

2. Informed sustainability design: BIM contributes to sustainable lifecycle decisions and processes because it leverages the ability of the entire construction value chain, thus optimising design decisions on complex issues such as energy efficiency.

3. Modelling standards: BIM promotes developing and adopting a wide range of standards and best practice guides, as evidenced by BIM adoption dynamics in Europe.

4. Increase in BIM uses for retrofit. There is an increasing trend for using BIM in large and small projects with the desired benefit of maximising energy efficiency and sustainable outcomes. Recognising the suitability of BIM for small retrofit projects is also critical, given the anticipated dynamic growth in the green retrofit market in existing domestic stock throughout Europe.

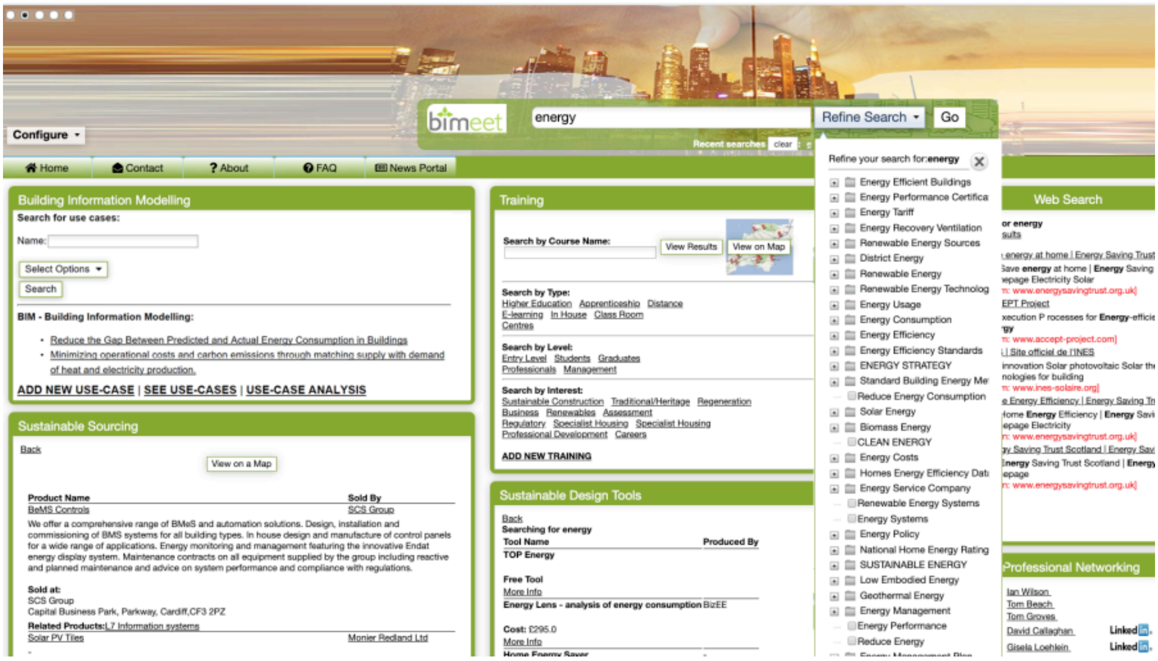


Figure 8: The energy-bim.com portal

5. Using BIM for building performance monitoring: There is increasing evidence of the value BIM tools during a project's operations and maintenance phase of a project, to reduce the endemic gap between the predicted and actual energy consumption in buildings.

6. Training support & communication tool: As BIM embraces building products and processes, it constitutes useful support for training and communicating the best practices for energy-efficient and high-quality construction, in particular, to on-site staff.

3.3.2 Conducting semantics analysis of the BIM Knowledgebase data sources

The ontology running in the background of the *energy-bim.com* portal has facilitated the mining and analysis phases of the BIM knowledge harvesting process [284]. The query and extension methods within the framework provide the primary use of an ontology to drive the search engine. First, the terms in the ontology are used when entering search terms (using the query method) to give keyword suggestions. Second, the relationships between words help users expand/limit their questions based on ontological suggestions.

The ontology service is built on the basis of semantic vectors to ensure the required level of knowledge for the platform is met. The ontology seeks to enhance and extend the contents and current domain requirements with additional concepts and aspects taken from (i) an engineering-specific knowledge repository and (ii) information structures that are the basis for calculations, simulations, and resources for monitoring compliance. The ontology uses algorithms such as "Term Frequency-Inverse Document Frequency (TF-IDF)" and metric clusters to detect related ontological concepts in and around a knowledge-base repository. The researcher measures the degree of significance (semantic) in more detail for each definition and facet of the text and the entire collected documentary repository. A process specifying the number of co-occurrences of concepts in the document was implemented to determine the importance associated with relationships between concepts, quantified at a numerical level. This clustering algorithm calculates the difference between two terms to measure their correlation factor.

The *energy-bim.com* platform has been developed based on preliminary consultations and user requirements, which have indicated that they require access to BIM knowledge and training resources. The researcher has incorporated scientific publications and regulations alongside Twitter data to create a scalable knowledge base environment accessible to users and actors from the construction industry. The platform has been used to support and deliver the analysis and the entire methodology of Stage III.

This stage of the methodology comprises a mixed content mining and analysis research method to determine BIM roles, skills, and competencies for the construction industry. The researcher carried out an incremental research process from content aggregation to data curation and analysis to extract knowledge related to BIM roles and associated skills for energy efficiency.

The researcher has carried out knowledge extraction on two data repositories on the *energy-bim.com* platform. Repository 1 identifies scientific publications, data and regulations, and repository

2 comprises a collection of tweets from social media sources.

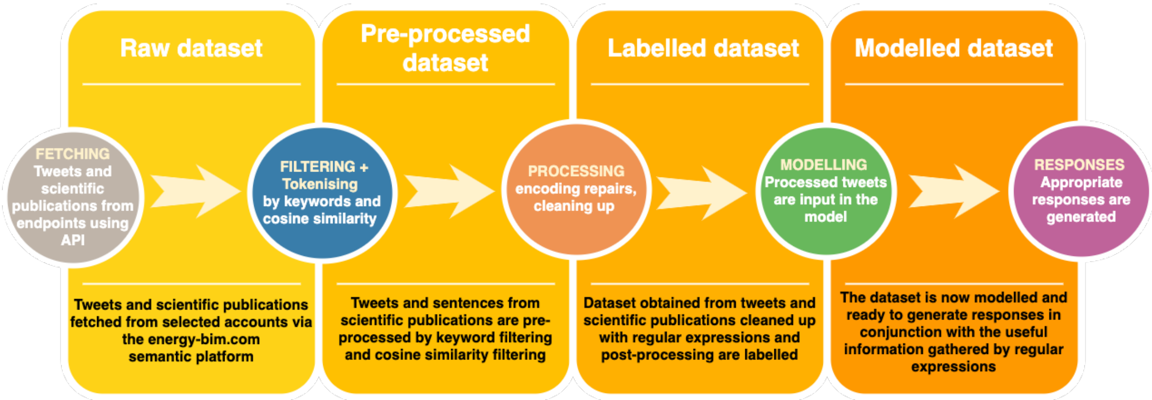


Figure 9: The steps of the process involved in the methodology

The steps covered by the methodology are broadly illustrated in Fig. 9. To confirm the roles and competencies required by this training process, a repository of scientific literature composed of 80 BIM publications and related analytics has been implemented using the above steps.

3.4 Stage II - Exploring the role of semantics in enhancing UX through a web-based User Interface for an Energy Management Platform

The navigation of the research journey leads us to the next research stage. Building on the foundational insights derived from the initial stage, this subchapter unveils the methodological approach for the second stage of the study, designed to address the second research question: *How to derive context from static artefacts (buildings) and dynamic artefacts using a Digital Twin?*.

In the sphere of thesis, context is understood as a comprehensive amalgamation of environmental, physical, and user-based information, all of which have the potential to significantly improve the user experience in web applications. When considered in the design and operation of web applications, context serves as a conduit for semantics, giving the software the ability to react intelligently and flexibly to changing user needs and environmental conditions [163].

This chapter delves deeper into contextualising semantics concerning an energy management platform. The exploration involves the integration of static artefacts, which in this case are the buildings, and dynamic artefacts facilitated through a Digital Twin. A Digital Twin, essentially a dynamic virtual replica of a physical object or system, allows us to collect real-time data and offer predictive insights, enhancing the ability to dynamically shape the user experience [348].

The primary focus is on deriving context from these static and dynamic artefacts, leveraging their potential to tailor the user interface of the energy management platform. In this case, the emphasis on context is driven by the understanding that an efficient and effective energy management platform must respond to static and dynamic changes in the environment [315].

In tune with the interpretative approach, this stage again embraces an iterative methodology involving a comprehensive review of the literature, gap analysis, implementation of the context-aware Digital Twin-based prototype, and continuous integration of user feedback. Through this process, the researcher sought to ensure that the energy management platform is not only technologically advanced but also highly attuned to the needs of its users.

To answer this research question, the researcher adopted an iterative approach to develop the Computational Urban Sustainability Platform (CUSP). The development of this platform was carried out in the context of Cardiff City Council and was then evaluated using a case study of a railway station in the UK.

3.4.1 CUSP - Motivations for developing an AI-powered ontology-based energy efficiency solution

As the exploration of CUSP has conveyed, semantic frameworks have significant potential to transform complex data structures into visually annotated, user-friendly interfaces within web applications for buildings and cities. The intricate modelling of these complex concepts via ontologies plays a pivotal role in enhancing user experience by promoting a deeper understanding of the system's data structure, fostering semantic interoperability, enabling personalisation, and supporting intelligent search capabilities.

The creation of CUSP is an essential part of the methodology. CUSP is a semantic and AI-enabled system that facilitates context inference from static and dynamic artefacts. This is an iterative process, which means that the platform undergoes continuous improvements and refinements based on user feedback, operational learning, and changing requirements.

The essence of CUSP is its ability to leverage the power of artificial intelligence to capture, analyse, and interpret the complex relationships and interactions among the building artefacts. It employs sophisticated machine learning algorithms and AI models to process large volumes of data in real time, identify patterns, extract pertinent insights, and predict future trends. These capabilities are essential to comprehend and enhance the user experience in BIM web applications.

In this context, semantics connects the AI model to real-world artefacts. The semantics of CUSP bridge the AI model's variables and the building's physical and virtual entities, allowing for the translation and mapping of abstract data into practical and actionable insights. The collected raw data can be deciphered by applying semantics and associating them with tangible and meaningful aspects of building and user interactions.

To build this bridge, a semantic model is created that contains a set of definitions and relationships for each variable in the AI model, connecting them to their corresponding real-world artefacts. The model enables CUSP to comprehend the collected data, map it to its respective real-world equivalent, and use it to generate a complete picture of user interactions and environmental context.

For example, a variable in the AI model could be related to the light intensity in a particular building section. The semantic model informs CUSP that this variable refers to the physical properties of light and their influence on user comfort and energy consumption. Then it can analyse these data along with other pertinent data, such as time of day, occupancy level, and user feedback, to optimise lighting conditions and improve the user experience.

This integration of AI and semantics enables CUSP to be adaptive, intelligent, and context-sensitive, resulting in a more personalised and enhanced web user experience for BIM applications. The feedback loop of the iterative methodology enables CUSP to learn from its operations and adapt its model and algorithms accordingly, fostering continuous learning and development.

3.4.2 An iterative approach - motivations

The decision to adopt an iterative methodology to develop a minimum viable product (MVP) of CUSP is based on several key considerations, all of which contribute to optimally addressing the second research question.

Increased flexibility: An iterative approach inherently fosters adaptability and flexibility, allowing the incorporation of knowledge gained and changes discovered during the development process into subsequent iterations [204]. Given the multifaceted nature of the derivation of the context from static and dynamic artefacts, the iterative method allows the MVP development process to adjust and evolve along with an increasingly sophisticated understanding of these intricate artefacts and their respective contexts.

User-Centred Design: The user is at the heart of the proposed research, focusing on improving the user experience on the Web for BIM web applications. An iterative development process promotes user testing and feedback during each cycle, allowing incorporation of these insights to fine-tune the features, functionality, and user interface of CUSP, thus ensuring a more user-centric design [260].

Risk Mitigation: Iterative development also helps mitigate risk. By dividing development into manageable parts, potential risks and challenges can be identified and addressed early, preventing large-scale problems later [54]. Therefore, an iterative approach significantly reduces the potential impact of failures.

Efficiency: The iterative method ensures that resources are allocated more efficiently, focusing on the most valuable functionalities at each stage [359]. This can lead to significant cost savings and more efficient use of time and resources compared to a traditional linear development approach.

Therefore, the iterative development of an MVP of CUSP aligns with the research goal of effectively deriving context from static and dynamic artefacts using a Digital Twin. Each iteration, marked by continuous learning and improvement, gradually progresses towards enhancing the web user experience for BIM web applications.

3.4.3 Prerequisites

A critical precursor to the development and refinement of CUSP is the meticulous gathering and understanding of site-specific requirements, a process rooted in the principles of requirement engineering (RE) [335]. Given that the case study focuses on a city within the UK, a thorough understanding of the architectural, functional, and operational nuances specific to the site, as well as the broader context of railway stations and similar infrastructures is needed. These requirements drive the customisation and adaptation of the AI models in CUSP and semantic relationships.

The next subsections describe each step of the CUSP User Interface Lifecycle, following an iterative approach.

3.4.4 Requirements Gathering and User Research

The initial step of obtaining requirements and user research sets the foundation for subsequent improvements. It ensures that the smart city visualiser aligns with the needs and expectations of users. Academic research by Nielsen [262] emphasises the importance of user-centred design and iterative usability testing throughout the development process to identify user needs and incorporate them into interface design. By involving stakeholders and conducting thorough user research, this step improves overall usability and user satisfaction of the visualiser [317].

The development process of a smart city visualiser, with Cardiff Council as the client, begins with comprehensive requirements gathering and user research. At this stage, the researcher collaborated closely with Cardiff Council to define the specific goals and objectives of the visualiser in line with the council's smart city initiatives. This approach involves understanding the strategic vision of the council for urban planning, citizen engagement, and data-driven decision-making.

The researcher conducted extensive user research, engaging with various stakeholders within the Cardiff Council, including city planners, architects, and other urban engineering specialists, eliciting the stakeholders' needs, expectations, and use cases through interviews, surveys, and workshops, striving to gain insights into their diverse requirements for a smart city visualiser. The researcher also examined existing data sources and infrastructure within the Council to identify potential integration points and data availability.

The limitations of this stage lie in the challenges associated with gathering comprehensive and accurate user feedback. Different stakeholders may have varying perspectives and priorities, requiring careful consideration and synthesis to form a cohesive understanding of the visualiser's requirements. Furthermore, accurately predicting all potential use cases at this early stage may prove challenging, as emerging needs and scenarios may arise during development.

However, the detailed requirements gathering and user research efforts lay the foundation for subsequent iterative steps in the development process. By engaging with Cardiff Council and stakeholders, the researcher ensured that the visualiser was tailored to meet the specific needs and expectations of the client, thereby aligning with Cardiff Council's strategic objectives for smart city development.

3.4.5 Conceptual Design and Wireframing

The next step in the process is the *conceptual design and wireframe* stage, which improves the interface by visually representing the structure and layout of the visualiser. As Camburn et al. [66]

emphasises, early stage design iterations are important to explore multiple design possibilities and make informed decisions about interface layout and organisation. This step improves usability and user experience by providing a visual blueprint for subsequent development stages.

The researcher refined the interface by creating wireframes based on user feedback, visual representations, and layout planning, precursors to the functional user interface. During this stage, the researcher worked closely with Cardiff Council and stakeholders to establish the visualiser's overall structure, layout, and flow, considering the specific needs and characteristics of Cardiff City.

Based on the requirements gathered in Step 1, the researcher collaborated with city planners and urban designers to conceptualise the visualiser's interface. For example, wireframes were created to show the placement of key components such as interactive maps, data overlays, navigation menus and search functions, taking into account the unique geography and landmarks of Cardiff City.

Inspired by the reality of Cardiff City, the researcher designed wireframes that highlighted key urban features and points of interest, such as iconic buildings like the Principality Stadium or landmarks like Cardiff Bay.

The researcher regularly met with Cardiff Council and relevant stakeholders throughout this stage, presenting the wireframes and gathering feedback to refine the conceptual design. For example, input from city planners and citizen representatives helped ensure that the visualiser emphasised sustainability initiatives, such as highlighting green spaces, promoting active transportation options, and visualising air quality data.

In addition, the wireframes also considered the needs of citizens and community groups by incorporating features that foster engagement and collaboration, including areas for public feedback, integration with social media platforms to share information, and tools for citizens to report issues or suggest improvements to city infrastructure or services.

The limitations in this step included the challenges associated with striking the right balance between functionality, usability, and aesthetics within the wireframes. Additionally, stakeholders' diverse perspectives and priorities required careful consideration and negotiation to ensure a comprehensive and inclusive representation of Cardiff City's smart city vision within the visualiser interface.

By iteratively refining the feedback-based wireframes and incorporating the unique aspects of Cardiff city, the conceptual design and wireframe stage laid the groundwork for the subsequent development of a visually appealing and user-friendly smart city visualiser tailored to the specific needs and characteristics of Cardiff city.

3.4.6 Visual Design

The visual design step focuses on improving the aesthetics and usability of the interface. As mentioned previously, Hassenzahl and Tractinsky [148] suggests that visual appeal positively influences users' perception of usability. Incorporating visual elements that align with the city branding and considering the principles of visual hierarchy and readability improve overall user experience and satisfaction

[149].

Step 3 of the development process involved designing the smart city visualiser for Cardiff Council, incorporating elements that align with the city's branding and enhance usability.

The researcher collaborated with Cardiff Council to develop a visual design that reflected the city's unique identity and visual language. Inspired by Cardiff's cultural heritage, the visual elements incorporated colours such as deep blue and vibrant green to evoke the city's maritime and natural surroundings.

Typography choices were made to enhance readability and maintain consistency with the existing communication materials of Cardiff Council. For instance, sans-serif fonts were used for headings and labels, while legible serif fonts were used for body text, ensuring clarity and accessibility of information within the visualiser.

Cardiff's architectural landmarks, such as Cardiff Castle and the Millennium Centre, inspired the iconography of the visual design. Icon sets were created to represent different categories of data, using simplified representations of notable city structures to provide visual cues that resonate with Cardiff residents and visitors.

The visual hierarchy of the interface design emphasised the importance of data visualisation. For example, Cardiff Council's environmental sustainability initiatives were visually highlighted through prominent graphs and charts showcasing energy consumption, waste management, and air quality data.

To promote inclusivity, the visual design incorporated features to accommodate various user groups, including those with visual impairments. High-contrast colour schemes, alternative text for images, and adjustable font sizes were implemented to ensure that the visualiser was accessible to all members of the Cardiff community.

The limitations in this step revolved around the challenge of striking a balance between aesthetics and usability. While visual appeal was important, it was crucial to maintain a user-friendly interface that facilitated easy comprehension and intuitive navigation.

The iterative design process allowed the researcher to address these limitations by continuously refining the visual design based on user feedback and usability testing. This approach resulted in an interface that harmoniously combined the unique aspects of Cardiff's identity with optimal user experience.

3.4.7 Development of the User Interface

The development of the user interface brings the visual design to life, allowing users to interact with the visualiser. This step improves the interface's functionality and navigation by implementing core functionalities and conducting usability testing. This aimed to enhance usability and user satisfaction by ensuring that the visualiser meets user needs and expectations [200].

Step 4 of the development process involves the actual implementation and development of the user interface for the smart city visualiser, incorporating low-fidelity OSM Buildings models for initial visualisation.

The researcher used front-end technologies, such as HTML, CSS, and various JavaScript frameworks, to implement the user interface for the smart city visualiser, ensuring compatibility and responsiveness across different devices and platforms.

To provide initial visual representations of buildings within Cardiff City, low-fidelity OSM Buildings models were integrated into the user interface. These models, generated from OpenStreetMap data, allowed users to navigate and explore the city's urban environment at a basic level.

Implementing core functionalities, including panning, zooming, and data overlays, enabled users to interact with the visualiser and access relevant information. For instance, users could explore different data layers, such as transportation networks, parks and green spaces, and points of interest within Cardiff City.

Throughout the development process, regular usability testing sessions were conducted to identify usability issues and gather user feedback. This feedback served as a valuable input to improve the user interface and the overall user experience.

The limitations in this step included the potential lack of detail and accuracy in the low-fidelity OSM Buildings models. However, their integration provided an initial foundation for testing and validating the functionality and navigation of the visualiser interface.

Iterative improvements were made based on user feedback and usability testing to address these limitations. For example, feedback from city planners and architects guided the enhancement of rendering techniques to improve the realism and accuracy of building models within the visualiser.

Collaboration with Cardiff Council and stakeholders also allowed for the identification of specific data integration needs, which led to the integration of additional data sets, such as real-time transportation information and environmental data, to provide users with more comprehensive and up-to-date information within the visualiser.

By continuously iterating on the implementation and gathering user feedback, the researcher was able to refine the user interface, ensuring a more accurate representation of Cardiff City and an improved user experience in navigating and exploring the smart city visualiser.

3.4.8 Data Integration and Conversion

Integrating high-resolution IFC models improves the visual fidelity and realism of the buildings within the visualiser. Academic research by Lurie and Mason [221] emphasises the importance of visual realism in data visualisations to improve user understanding and engagement. This step aimed to improve the user experience and support data-driven decision making processes by providing more accurate representations of buildings [153].

Step 5 of the development process involves data integration and conversion, including the acquisition of high-resolution IFC models of buildings and their integration into the smart city visualiser for Cardiff City.

The researcher collaborated with relevant data providers and stakeholders to acquire high-resolution IFC models of buildings in Cardiff City. These models contained detailed geometric and semantic information, including architectural elements and building attributes. Tools and technologies were developed or used to convert acquired IFC models into a format compatible with the smart city visualiser. This conversion process ensured a seamless integration of the high-resolution models into the existing interface.

The integration of high-resolution IFC models replaced the previously used low-fidelity OSM Buildings models within the visualiser. This enhancement resulted in more accurate and detailed representations of the buildings in Cardiff City, showcasing architectural intricacies, textures, and precise geometries.

For example, the integration of high-resolution IFC models allowed users of the visualiser to explore iconic buildings such as Cardiff Castle and the Wales Millennium Centre with a higher level of fidelity and realism. Detailed representations of these structures gave users a more immersive and engaging experience.

The data integration pipeline was established to import and process high-resolution IFC models within the smart city visualiser. This pipeline ensured that the models were updated regularly, reflecting any changes or additions to the buildings in Cardiff city, and maintaining the accuracy and currency of the visual representations.

Limitations in this step may include the availability and accessibility of high-resolution IFC models for all buildings in Cardiff City. Acquiring comprehensive and up-to-date data for the entire city can be a significant challenge, requiring collaboration with data providers and ongoing efforts to expand the coverage of the data set.

However, integration significantly improved the fidelity and realism of visualisations within the smart city visualiser. Users could explore and analyse the built environment of Cardiff City with greater precision, supporting urban planning, architectural visualisation, and decision-making processes.

By continuously refining the data integration process and ensuring the compatibility of high-resolution IFC models, the smart city visualiser for Cardiff City evolved to provide more accurate, detailed and up-to-date representations of the buildings, enhancing the overall visual experience for users.

Integrating high-resolution IFC models in Step 5 of the development process for the smart city visualiser brings about substantial positive differences compared to the low-fidelity OSM Buildings models used in step 4. These differences result in enhanced usability of the user interface and improved user experience, as supported by academic research in Human-Computer Interaction (HCI).

High-resolution IFC models give users a more accurate and detailed representation of the buildings within the visualiser interface. Research by Marwecki et al. [234] emphasises the importance of visual realism in interfaces, highlighting that realistic representations enhance user engagement and comprehension. The high-resolution models offer a more immersive and realistic experience for users exploring Cardiff City's built environment by showcasing architectural intricacies, textures, and precise geometries. This improvement in visual fidelity contributes to a greater sense of presence. It enables users to better understand the urban landscape, facilitating architectural visualisation and urban planning activities [175].

Furthermore, the integration of high-resolution IFC models positively impacts the user experience by providing a more engaging and informative interface. The research carried out by Bargas-Avila and Hornbaek [30] underscores the importance of information richness in interfaces, stating that more detailed representations enable users to better interpret and comprehend complex information. The high-resolution models offer users a wealth of visual data, allowing for detailed examination and analysis of the buildings. This approach supported decision-making processes for urban design, heritage preservation, and infrastructure development [175].

Users can explore iconic structures in Cardiff City, such as Cardiff Castle or the Wales Millennium Centre, with a higher level of fidelity and realism. This approach contributed to a more satisfying and captivating user experience, as validated by studies on the positive relationship between visual appeal and user satisfaction [351].

In general, integrating high-resolution IFC models in Step 5 significantly improves the usability and user experience of the smart city visualiser. By providing a more accurate and detailed representation of buildings, the interface becomes more engaging, informative, and realistic, supporting architectural visualisation, urban planning, and decision-making processes. The positive differences resulting from this improvement align with academic research in the field of HCI, highlighting the importance of visual realism and information richness in interfaces to enhance user engagement and satisfaction.

The positive visualisation differences also lie within the context of chart data visualisation and the ability of a facilities manager to easily monitor key metrics of building stock.

Research by Heer and Shneiderman [153] highlights the importance of effective data visualisation to facilitate decision making and understanding complex information. The high-resolution IFC models enable the visualiser interface to display detailed charts and graphs representing key metrics of the building stock in Cardiff City. For example, a facility manager can easily access visual representations of energy consumption, water usage, or occupancy rates. This improved visualisation of the data supports the faster identification of trends, anomalies, and areas that require attention, ultimately helping to optimise facility management and resources [96].

High-resolution models also improve user experience by providing an intuitive and accessible interface for facility managers. Norman [266] emphasises the importance of the learnability and ease of use of an interface in reducing cognitive load and promoting user satisfaction. With high-resolution models integrated into the visualiser, facilities managers can effortlessly navigate and interact with the 3D representation of the building stock. They can zoom in, rotate, and examine individual buildings or compare metrics between different buildings, fostering a comprehensive understanding of the data and facilitating data-driven decision-making.

The positive differences resulting from the integration of high-resolution IFC models, as compared to low-fidelity OSM Buildings models, align with the principles of effective data visualisation and user-centred design advocated in the field of Human-Computer Interaction. By providing detailed chart data visualisations and facilitating easy monitoring of key metrics, the visualiser interface enhances usability and user experience, empowering facility managers in their decision-making processes and optimising building stock management.

3.4.9 Continuous Improvement and User Feedback

The continuous improvement process based on user feedback ensures that the smart city visualiser evolves to meet the changing needs of users. Research by Nielsen [260] highlights the importance of incorporating iterative user feedback throughout the development process to identify usability issues and make iterative improvements. This step improves usability and user experience by addressing usability issues, optimising performance, and implementing new features based on user needs and preferences [106].

In this case, Step 6 of the development process involves refining the user interface in response to user feedback and deciding to progress from meeting core functional requirements to the enhancement stage.

Throughout the development process, regular usability testing sessions were conducted with representatives of Cardiff Council, facilities managers, and other stakeholders to evaluate the user interface and gather feedback. Usability testing included tasks such as navigating the visualiser, accessing specific building metrics, and interpreting data visualisations.

Based on feedback received, iterative improvements were made to address usability issues and enhance user experience. For example, during usability testing sessions, facilities managers expressed the need for an intuitive dashboard that would provide a holistic overview of key metrics for the entire building stock in Cardiff City.

Once a point where the core functional requirements of the visualiser were met effectively was reached, a critical decision point was reached. This decision was based on a rigorous evaluation of user feedback and assessment against predefined success criteria. The visualiser was determined to adequately fulfil the fundamental goals of providing real-time data visualisation, facilitating facility management, and supporting data-driven decision making for Cardiff City.

The decision to progress to the enhancement stage was driven by several factors, including user satisfaction with the core functionalities, successful completion of usability testing tasks, and alignment with the established objectives of Cardiff Council's smart city initiatives. Feedback from facilities managers and other stakeholders indicated that the core functional requirements were met, and the visualiser proved to be a valuable tool to monitor building metrics, identify trends, and support proactive facilities management strategies.

At this stage, the researcher and Cardiff Council agreed to prioritise additional enhancements to improve the user interface further and expand its capabilities. Enhancements could include advanced data filtering options, predictive analytics, or integration with the Internet of Things (IoT) devices for visualisation of sensor data in real time.

The decision to progress to the enhancement stage was driven by a thorough assessment of the visualiser's core functionality, user feedback, and alignment with the goals of Cardiff Council's smart city vision. This milestone marked a crucial transition from meeting essential requirements to further refining and expanding the visualiser's features and capabilities to deliver a more robust and user-centric smart city solution for Cardiff.

Step 6 of the development process involves refining the user interface in response to user feedback and deciding to progress from meeting core functional requirements to the enhancement stage.

Throughout the development process, regular usability testing sessions were conducted with representatives of Cardiff Council, facilities managers, and other stakeholders to evaluate the user interface and gather feedback. Usability testing included navigating the visualiser, accessing specific building metrics, and interpreting data visualisations.

Based on feedback received, iterative improvements were made to address usability issues and enhance user experience. For example, during usability testing sessions, facilities managers expressed the need for an intuitive dashboard that would provide a holistic overview of key metrics for the entire building stock in Cardiff City.

Once a point where the core functional requirements of the visualiser were met effectively was reached, a critical decision point was reached. This decision was based on a rigorous evaluation of user feedback and assessment against predefined success criteria. The visualiser was determined to adequately fulfil the fundamental goals of providing real-time data visualisation, facilitating facility management, and supporting data-driven decision making for Cardiff City.

The decision to progress to the enhancement stage was driven by several factors, including user satisfaction with the core functionalities, successful completion of usability testing tasks, and alignment with the established objectives of Cardiff Council's smart city initiatives. Feedback from facilities managers and other stakeholders indicated that the core functional requirements were met, and the visualiser proved to be a valuable tool to monitor building metrics, identify trends, and support proactive facilities management strategies.

At this stage, the researcher and Cardiff Council agreed to prioritise additional enhancements to improve the user interface further and expand its capabilities. Enhancements could include advanced data filtering options, predictive analytics, or integration with the Internet of Things (IoT) devices for visualisation of sensor data in real time.

3.4.10 Continuous deployment and improvement

The continuous deployment and improvement stage plays a crucial role in the smart city visualiser development process, ensuring ongoing enhancements and optimisations based on user feedback and data analysis. This stage is essential to maintain a user-centric approach and deliver a successful and evolving smart city solution.

The research by Malakhatka et al. [225] emphasises the importance of continuous improvement in software development, stating that it allows the refinement of user interfaces and the addition of new features to better meet user needs. Additionally, Mishra et al. [247] highlights the benefits of continuous deployment in mobile app development, indicating that it enables quick responses to user feedback and allows timely updates and improvements.

The continuous improvement process enables the researcher to gather insights from user interactions, analyse user behaviour, and identify areas for improvement and optimisation. This approach aligns with user-centred design principles, highlighting the importance of iterative design and user feedback [266].

By continuously refining the user interface, addressing usability issues, and incorporating new technologies and data sources, the smart city visualiser can remain relevant, up-to-date, and aligned with the changing needs of Cardiff City and its stakeholders. This iterative approach to continuous improvement ensures that the visualiser evolves alongside technological advancements and user requirements, ultimately delivering a more valuable and user-centric smart city solution for Cardiff.

Following completion of the user interface development, the smart city visualiser was deployed in a production environment to ensure its availability to the target users in Cardiff City. This involved considerations such as hosting infrastructure, scalability, and security measures to safeguard the integrity of the visualiser and user data.

After deployment, the researcher initiated a continuous improvement process to enhance the smart city visualiser based on user engagement and feedback. The researcher monitored user interactions and collected feedback from various stakeholders, including Cardiff Council, facilities managers and citizens, to identify areas of improvement and optimisation.

The analytics and user data were analysed to gain insight into user behaviour and usage patterns within the visualiser. For example, data on the most frequently accessed building metrics, popular data visualisation features, or commonly performed tasks were examined to identify areas of high user interest and potential pain points.

Based on user feedback and data analysis, iterative updates and refinements were made to the visualiser's interface and functionality, which included addressing usability issues, optimising performance and responsiveness, and implementing additional features to improve the overall user experience.

For example, in response to feedback from facilities managers, new features were introduced to enable customisable data visualisation dashboards, allowing them to focus on specific building metrics most relevant to their responsibilities and priorities.

The continuous improvement process also involved the collaboration with third-party partners and the integration of new technologies and data sources. For example, partnerships with utility providers allowed real-time integration of energy consumption data, allowing facility managers to monitor and optimise energy use in city buildings in Cardiff City.

The iterative nature of the continuous improvement process ensured that the smart city visualiser evolved alongside the changing needs of Cardiff City and advances in technology. This iterative approach, guided by user feedback and data-driven insights, contributed to the ongoing enhancement and optimisation of the visualiser, ultimately delivering a more valuable and user-centric smart city solution for Cardiff.

3.5 Stage III - Scaling up semantics at the city level with citizens as active sensors

Having addressed the second research question in the previous chapter, which focused on deriving context from static and dynamic artefacts using a Digital Twin, the researcher now embarks on the third stage of the methodological approach. This stage focuses on answering the following research question: *How can we rely on this diverse (i.e. tangible and intangible) dynamic context to enhance a user's experience (e.g., a stadium, city planner, city manager) in a wide range of applications applied in the built environment?*

The scope of this research question expands the inquiry beyond specific applications and seeks to understand the broader implications of utilising diverse, dynamic contexts in the built environment. The dynamic context encompasses tangible elements, such as physical structures and infrastructures, and intangible factors, such as social, cultural, and environmental aspects, that influence user experiences.

The researcher aims to uncover the mechanisms, methodologies, and strategies that can effectively leverage the dynamic context to cater to the needs and expectations of various users, including stakeholders like stadium administrators, city planners, and city managers.

With social media adoption expanding exponentially and accelerating by the widespread use of mobile devices, users increasingly react in real time to events occurring in their immediate surroundings and beyond, providing time-critical and potentially actionable information in a smart city context. With population growth and technological advancements, including the deployment of 5G, the potential for citizens to act as "active" social sensors and actuators" has become increasingly promising. In 2021, social networks were well established, with more than 75% Americans using at least one platform, according to Faelens et al. [116]. Numerous technical factors contribute to this trend, such as the increasing prevalence of global Internet availability. According to Russo [310], in 2018, approximately 99% of European households had access to a 4G connection, compared to 20% in 2012. Another contributing factor is the widespread use of social networks, which governments are increasingly using to communicate directly with citizens [376]. These breakthroughs enable more and more "always-on" citizens to react in real-time to events, regardless of location, while openly reporting their internal thought processes to the public.

As social media applications such as Twitter and Facebook exceeded email as the primary communication channel for most users [224], they accumulated ever increasing information of interest for data analysts. One of the leading social networks for short-form messages, Twitter has an average daily usage rate of 500 million user posts. Large proportions of the population of developed countries live in conglomerated urban areas. As described by Malche et al. [226], among other functions, smart cities monitor various parameters such as energy, water quality, and transportation flow while detecting and preventing crime and organising appropriate responses proportionate to identified risks and events.

In an increasingly urbanised and digital world, cities are exploring a wide range of governance models, informed by decision support systems that leverage (near) real-time information, including social media sources, to improve their sustainability and resilience [189]. As such, the Smart City paradigm seeks to improve the standard of living of its residents. To determine whether smart city development is consistent with the sustainable development goals, a new methodological approach has gained traction. Girardi and Temporelli [129] defined *smartainability* as an approach that aims to evaluate, using qualitative and quantitative measures, how sustainable smart cities become due to the implementation of smart technologies and infrastructure. The author implemented this strategy on the *Expo Milano 2015* site, demonstrating the ability to provide decision makers with valuable information on the following advantages generated by the deployment of smart solutions at the city level: (i) benefits can be measured, (ii) indicators can be assessed prior to implementing technologies or solutions, and (iii) benefits could be easily linked with the associated smart technologies deployed for a more targeted assessment. This research stage devises a viable approach to achieve and maintain *smartability*. This paradigm aims to enable smart city managers, such as stakeholders and local authorities, to make more informed decisions around detected events, prioritised under the analysis of their associated environmental factors and citizen satisfaction.

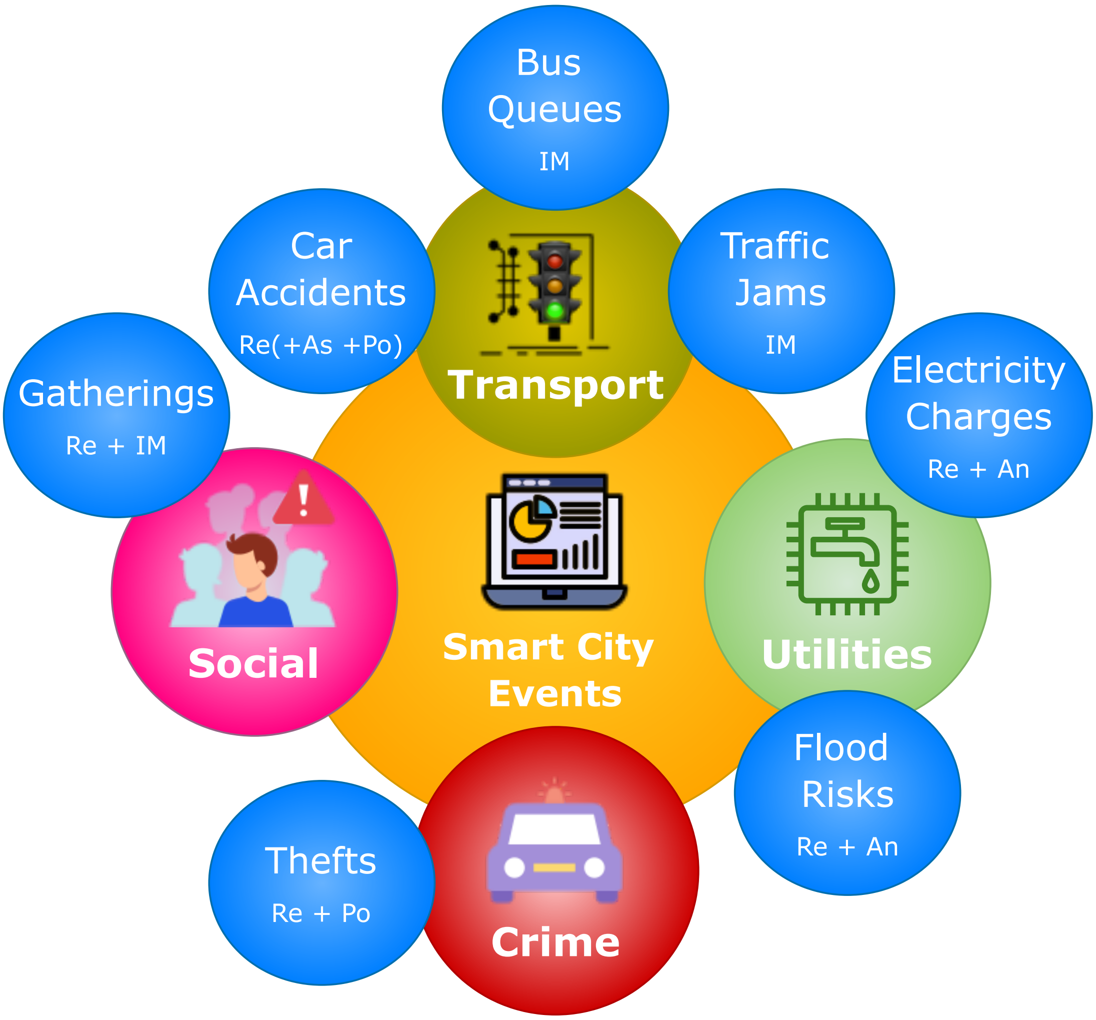
3.5.1 A custom taxonomy for smart city event detection and citizen satisfaction analysis

A *taxonomy* built upon a semantic-based risk model is presented. The researcher also proposes a method to conduct a citizen satisfaction analysis, which would enable smart city managers to better gauge the well-being of residents and transients in their areas of interest, while also evaluating the importance of particular detected events more objectively. A citizen satisfaction metric is determined on the basis of each message's positive or negative valence. In addition, the researcher uncovers and validates relationships between co-occurring events, citizen satisfaction, and environmental factors. Additionally, the researcher validates the relationships using MRA statistical models.

Each message published on Twitter features valuable meta-data, such as originating location and other user-defined attributes and content (images and videos). As a result, manually processing and decoding voluminous data streams becomes increasingly difficult, if not impossible. Notwithstanding this barrier, the researcher can gear the Twitter limitations to the advantage: the highly condensed and standardised format of the messages imposed by the platform, limited to 280 characters each, facilitates the deployment of automated models for aggregating, processing, and decoding "events" of interest within the data, with relatively low computational costs.

From a technical perspective, an event represents a real-time occurrence identified within a message that, based on the determination of machine learning NLP techniques, features a semantic similarity above a certain threshold defined in the hyperparameters. On the contrary, from an ontological perspective, an event acts as a stimulus. It has the consequence of disturbing the modus operandi of a city, taking place either in a tangible form (*thefts*, *car accidents*, *social gatherings*) or in an intangible form (*positive* or *negative* sentiments).

Figure 10: Smart City Taxonomy example and appropriate local authority responses: IM = Interactive map, Re = Report, As = Assistance, An = Analyse, Po = Police, Fi = Firefighter, Te = Technician



An efficient approach to addressing risk concerns in a smart city is to react promptly to emerging threats. Using the work of Coburn et al. [83] as a starting point, the researcher has derived a general risk *taxonomy* illustrated in Figure 10. Although mentioned and modelled by the original author, events warranting a deeper level of analysis, such as wars, nuclear, space, and economic threats, have been omitted, as they do not form part of the remit of this research. Although not exhaustive, the *taxonomy* provides a strong starting point for event detection in a smart city context. At this stage, the researcher considered seven categories of events (highlighted within the blue circles) as a starting point. In the taxonomy highlighted above, the researcher has also proposed appropriate responses to each type of event.

3.5.1.1 Citizen satisfaction and environmental factor analysis

As some decisions taken in a smart city context could directly and significantly affect the livelihoods of the citizens, they should be subject to a periodic review process. Using messages from a social media channel such as Twitter gives us an insight into the publicly exposed thought processes of residents and transients in a particular area. By analysing *positive* or *negative* inclinations, specific speech patterns and words acquire a *negative* or *positive* value, aiding in the prediction. Inspired by Lamba and Madhusudhan [203], an extensive collection of tweets was obtained and used to ensure a highly accurate category prediction. The extensive vocabulary variations allowed us to account for the subtleties of language and jargon that express human sentiments.

To analyse environmental factors, the researcher uses daily median values for *temperature*, *humidity* and *precipitation* fetched from *Meteoblue* [68]. This platform compiles data from various local weather stations and national weather services and incorporates it into model simulations to provide high-precision data at high spatial resolution.

3.5.1.2 A case study for a Smart City

Among several strategic research initiatives, the smart city and community programmes funded by the European Commission are of particular significance, including, but are not limited to: Smarter Together [250], Making city [80], CityxChange [349], ATELIER [27, 71, 272], MAtchUP [88, 102].

A heuristic analysis of these projects, coupled with the recurring event types and use cases modelled as part of the *Taxonomy*, shortlisted several relationships and scenarios of particular significance for smart cities:

- The interplay between *traffic events* and *environmental factors*
- The influence of *faulty infrastructures*, such as *electricity charges*, on *traffic events*
- Relationships between *thefts*, *social events* and *environmental factors*
- Links between *citizen satisfaction* and *gatherings*, coupled with *queues* and *electricity charges*
- Correlations between citizen satisfaction and weather elements (*temperature*, *precipitation*, and *humidity*)

These scenarios resemble real-world use cases in a smart city context, in line with recurrent focal topics highlighted within the European Commission-funded projects mentioned above. The researcher has associated each scenario with a dependent variable and multiple independent variables derived from the risk types modelled in the *taxonomy*.

Cardiff City Council has an ambitious agenda for the adoption and implementation of smart city technologies through its released "Smart City Roadmap" [87], which outlines proposals for the introduction of smart street lighting, smart parking, smart transportation, and a smart environment. Most of these initiatives aim to address some of the identified challenges: increased demand for public services and energy, uncertain economic conditions, and increased pressure on the natural environment. The tailored smart city research can assist Cardiff Council and other local authorities in developing an Open Source Intelligence (OSINT) framework to make evidence-based decisions considering historical and real-time events. This approach enables Cardiff City to maintain a competitive edge in the smart city revolution currently occurring in the UK and to remain a secure and highly available hub for sustainable development.

In a real-world scenario, a smart city manager would monitor the "urban pulse" using an interactive dashboard that evaluates different types of events at the city level. Events are detected from continuous messages and classified into a risk category. The events are geo-localised wherever possible and can be validated through manual review through cross-checking with third-party data sources, such as intelligence reports from local authorities and public service providers. When significant correlations are detected with citizen satisfaction or environmental factors in areas of interest, they are displayed automatically alongside the detected events, with suggestions on the significance of the associated event in terms of severity and urgency. The city manager would then conduct a risk assessment and assign proportionate resources, according to an appropriate intervention plan, according to the information collected in Table 2 below:

Table 2: Detected risk events and appropriate actions suggested on a dashboard to a Smart City manager

Event and area	***Suggested Action(s)***	***Priority***	***Background Information***
Fire, Roath Park	Dispatch a Fire Department unit to neutralise the fire	HIGH	• local residents indicated *negative* emotional predilection of *high intensity* • environmental factors (*high temperature*, *low humidity*, *low precipitation*) indicated a *high risk*
Bus queue, Cardiff City Centre	Reroute crowded bus services to increase traffic fluidity	MEDIUM	• local residents indicated *negative* emotional predilection of *moderate* intensity • environmental factors (*low humidity* and *low precipitation*) indicated a *below-average risk* to road safety in the short term • no recent *car accidents* or *traffic jams* detected in the area

Flood risk, Cardiff Bay	Actuate barrage movements to release water into the Celtic Sea	LOW	• local residents indicated *negative* emotional predilection of *low intensity* • environmental factors (*high temperature* and *low precipitation*) indicated *low risk* in the short term

3.5.1.3 Body of knowledge contributions

As it can be observed from the narrative illustrated above, this stage of the research covers the following areas of interest:

- the impact of the social factor on the resilience of cities
- monitoring and analysing actions and behaviours of citizens within urban communities
- smart city governance
- smart city and societies management optimisation through decision support systems for trade-off and uncertainty analysis
- case studies for Big Data, machine learning, and AI (Artificial Intelligence).

At this stage of the research, the researcher addresses a significant gap in the existing knowledge base on smart cities and sustainable communities. This is achieved by providing a comprehensive context, underpinned by primary sources, to a broad spectrum of sociotechnical and environmental phenomena occurring within urban landscapes. The key contributions of this research are as follows:

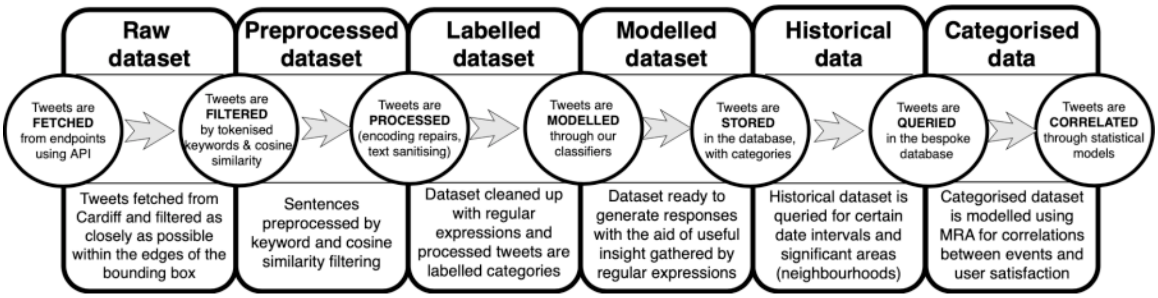
Firstly, the research innovatively positions citizens as integral sensors within the smart city paradigm, focusing on their role in reporting risks identified through a bespoke risk taxonomy. This approach uses natural language processing (NLP) techniques to effectively interpret the vast influx of data from social media platforms, thereby facilitating the detection of incidents as reported in real time by eyewitnesses.

Furthermore, the research methodically identifies patterns and trends related to the occurrence of risk events, levels of citizen satisfaction, and various environmental factors on the urban scale. An integral part of this analysis involves examining the interrelationships between citizen satisfaction, environmental variables, and risk events. These relationships are subsequently subjected to rigorous statistical validation through the application of Multiple Regression Analysis.

Through these methodological approaches, the researcher contributes significantly to the understanding of dynamic interactions within smart urban environments, offering valuable insights into the optimisation of urban living experiences.

With a clear delimitation of the scope of the analysis defined in Section 3.5.1, we could then outline the approach to data sourcing, constructing the source dataset and preprocessing. This section presents a high-level overview of the application architecture, together with the selected language classification models, based on a comparative preemptive evaluation and a baseline definition for event detection.

Figure 11: A high-level overview of the steps implemented as part of the proposed approach



To meet the objective of this research stage, the researcher adopts a multistage methodology, as presented in Figure 11. The process starts with a data pipeline based on raw data. Tweets are fetched from the Cardiff City area and filtered as closely as possible within the geographical delimitation. The sentences of the messages retrieved are then preprocessed and filtered by keywords and *cosine similarity*. The text is then sanitised (curated) using regular expressions and assigning categories with the help of the *Classifier*. The labelled dataset, combined with the useful insight gathered by regular expressions, is now modelled and ready to generate responses, formulating the historical dataset and splitting it into areas (neighbourhoods) and time intervals. Finally, the categorised data set is modelled using MRA for correlations between cooccurring events, citizen satisfaction (*positive/negative texts*) and environmental factors (*temperature*, *precipitation*, and *humidity*).

This chapter has presented a summary of the methodological approaches intended to be applied for each stage of the Research Design presented in Section 3.2. The following three chapters will provide a detailed account of the deployment / application of the above methodology in the context of BIM web applications, smart cities, and social media mining.

The methodology described within this chapter laid the groundwork for the technical implementation and practical applications of research. The following three chapters provide a comprehensive account of the hands-on work undertaken to bring these methodologies to fruition and demonstrate their practical value.

In the next work chapters, the researcher delves into the technical aspects of implementing the proposed solutions, showcasing how the research findings and methodological approaches discussed so far have been transformed into tangible applications. The emphasis shifts from theoretical frameworks to the actual execution of projects, encompassing the development of software prototypes, data collection and analysis, and real-world deployments.

Each work chapter corresponds to a distinct phase of the research process, outlining the specific goals, methods, and outcomes associated with that phase. These chapters encapsulate the practical implementation of the research design interventions mentioned earlier, such as enhancing existing web applications, developing prototypes, and iteratively integrating user feedback.

Furthermore, these chapters present the practical applications of the methodologies within the broader context of the built environment. They demonstrate how the proposed solutions can be applied in real-world scenarios, addressing the needs of various stakeholders, such as building professionals, energy managers, city planners and citizens.

Throughout these chapters of work, the researcher provides detailed accounts of the technical challenges encountered, strategies to overcome them, and the results achieved. The researcher emphasises the practical relevance of the research, highlighting the impact of the work on enhancing user experiences, improving energy efficiency, and facilitating decision-making in the built environment.

By documenting the technical implementation and practical applications of the methodologies, these work chapters contribute to the overall objective of the thesis - to enhance the user experience of BIM web applications through the utilisation of semantics. They provide a comprehensive narrative that bridges the gap between theory and practice, showcasing the tangible results of the research and underscoring the potential for real-world impact.

Within the subsequent work chapters, the researcher delves into the technical details and presents the results of his practical implementation, shedding light on the transformative potential of the research in the realm of BIM web applications and the broader built environment.

The subsequent chapters categorise the studies and findings of this research.

4 Enhancing learning and training through semantics

This chapter focuses on an in-depth analysis of the primary sources of evidence to identify the roles and associated skills necessary to improve the implementation of BIM for energy efficiency training in the construction market. This involved the analysis of a large corpus of documents and Twitter data using text mining techniques to extract BIM roles and skills and infer their associations and links.

An approach was proposed to address the first research question consisting of four phases from data set collection and filtering to semantisation and knowledge extraction. The approach is delivered through the use of a semantic web platform which automates the update of the roles and skills required to reflect the changes in adoption and deployment of BIM in the industry.

This chapter aims to address the first research question: *How could an existing web application be enhanced using context in a way that maximises User Experience?*

This question will be answered by demonstrating how building information modelling (BIM) can be used to inform the skills and roles required for the digital transformation of the construction industry in order to maximise the industry's efficiency. User experience is captured through qualitative sources on social media.

As mentioned in the Methodology section, this involves reliance on a publication repository created after searching popular scientific portals such as Google Scholar and Web of Science using specific keywords to filter relevant results to BIM for energy efficiency.

The next sections elaborate in each stage of the process, as described in Fig. 9 of the Methodology section.

4.1 Upskilling work in the construction industry

The progression towards smarter cities underscores the crucial role of the construction industry in shaping the future of urban landscapes. Emerging responsibilities not only revolve around infrastructure creation but also extend to the promotion of sustainable urban environments. With an increasing emphasis on climate change mitigation and energy conservation, the construction workforce is tasked with the assimilation of new areas of knowledge and technical capabilities, in particular, the nuances of BIM and semantics [394, 386, 97, 47].

This research stage asserts that upskilling the construction workforce and integrating advanced semantic-rich BIM processes are central to addressing contemporary urbanisation challenges. It further argues that adopting these innovative technologies will form the cornerstone of creating digital twins of buildings and cities, serving as dynamic entities that reflect real-time changes and enable predictive analyses [379, 395, 174].

4.2 Introduction to the BIMEET platform

A comprehensive literature search based on several matching criteria for the title, abstract, and keywords for papers was first conducted through the Scopus search engine for scholarly publications. The researcher has used several keywords to conduct these, such as "building information modelling", "energy efficiency", "role", "skill", "training", "education" and "competencies". This review identified articles published between 2009 and 2019, which have been imported into the *energy-bim.com* platform as a knowledge repository and used in the analysis phase. The fairly recent time interval is justified by the fact that BIM for energy efficiency is a novelty in technological advancements.

The end goal of the researcher is to gather a better understanding of the roles, responsibilities, and actions in the construction industry, so that they can be used as a filter mechanism when users search the web, helped by the semantic technologies behind *energy-education.com*. For example, if a user searches for "roles", they will only see results for "roles" rather than "responsibilities" or "actions". When users search, they are looking for a particular piece of information. Search results would be readily available and tailored to their needs. This also represents a factor that enhances the web user experience of the platform.

The social media repository was created using the *energy-bim.com* platform, which fetched a corpus of 40 million tweets from a selection of Twitter accounts belonging to friends and followers of the most prominent users of the *energy-bim.com* platform (illustrated in Table 3). In the filtering stage, the tweets were filtered using regular expressions and SQL queries, thus obtaining a corpus of 60,000 selected tweets for further analysis. The association & importance stage determines the pairs of roles and skills while computing an importance score for each role and skill. At the final stage (correlation), the TF-IDF and Metric Cluster methods are applied to establish concept dependencies for roles and skills. This has resulted in an ontology of concepts of roles and skills, together with their associated correlation factors. Below, the researcher explains the entire data collection and analysis process supported via the *energy-bim.com* platform.

The BIMEET /INSTRUCT project portal (energy-education.com), illustrated in the figure above, brings together nine partners around BIM technology as a key digital support for the energy efficiency of the built environment. Participants include the Luxembourg Institute of Science and Technology (LIST), Cardiff University, Centre Scientifique et Technique du Bâtiment (CSTB), Building Research Establishment Ltd (BRE), La plateforme Formation & Évaluation de l'INES, VTT Technical Research Centre of Finland Ltd, House of Training, Metropolia University of Applied Sciences and Centre for Renewable Energy Sources (CRES).

The researcher has adapted and re-developed a web-based platform solution that provides integrated access to building information modelling (BIM) resources to support the methodology and create a dynamic community to capture BIM training requirements. The platform is an open, scalable and polymorphic context-based solution with modules that enable serendipitous BIM information and knowledge discovery using a symbiosis of technologies such as semantic web and social networking.

This web-based platform provides integrated access to BIM Resources in the form of interactive, dynamic, and user-orientated services that take full advantage of the latest technology advances. The platform is an open, scalable, polymorphic context-based solution with modules that allow the use of a symbiosis of technology to unlock information and knowledge [159].

This platform has helped in the process of BIM training requirements for energy efficiency, but also aims to address the key issue of dissemination of knowledge and participation of stakeholders in BIM practices and construction. The objective is to identify gaps and requirements as an initial phase, but also to support the project implementation phase by providing construction professionals with the necessary training to provide effective BIM expertise in energy-efficient and low-carbon solutions, while also enabling them to use the latest innovative practices and regulations.

The results of the consultations and the exploration of key perceived barriers identified and identified the need to create impact circles that connect building professionals, energy administrations, and people with shared sustainability expertise to address a number of issues, including:

- lack of sharing, exploitation and re-use of isolated sustainable practices and principles
- lack of knowledge and understanding among key building stakeholders and end-users
- lack of easy access to organised sustainability information and knowledge
- an unmistakable connection between sustainability values and existing construction regulations and standards
- confusion as to the expense of sustainable solutions/technologies

Various actors have created different libraries, coordinated activities, etc., to better understand the evolving construction environment, but with little to no coordination to the knowledge of what each other is doing. The researcher has recognised from the results that a socio-technical "information approach" would resolve the issues described above. This led to the introduction of a one-stop-shop web-based platform. The key themes that arise can be summarised as the following guidelines for the proposed solution in terms of information specifications and functionality/services.

The researcher delineates the specific information requirements necessary for the successful implementation of the energy-bim.com platform. These requirements are essential to facilitate a comprehensive understanding and effective management of energy-related data and concepts. The following key areas are identified:

Categorisation of Information: A systematic approach to categorising diverse information types and topics is paramount. This includes a broad spectrum of topics such as laws, legislation, analytical methodologies, and innovation. Such categorisation is essential for structuring and simplifying the vast array of information, making it more accessible and comprehensible.

Focusing on Best Practice: Prior to the advent of the energy-bim.com platform, much of the relevant information existed at a very abstract or "high level". Therefore, there is a critical need to refine this information, focusing on the identification and dissemination of best practices. This approach aims to provide more practical and actionable insights.

Information and Knowledge Management and Sharing: Having access to information is insufficient; there is a necessity to transform this information into actionable knowledge. This transformation process requires advanced content management strategies and the contextualisation of new concepts. By doing so, the platform can bridge the gap between raw data and practical knowledge.

Providing Avenues for Marketing and Connectivity The platform must serve as a conduit for marketing and connectivity, especially for the supply chain and information related to the products. This aspect is crucial for stakeholders to access relevant market insights and forge connections within the industry.

These information requirements are integral to the development and effectiveness of the energy-bim.com platform, ensuring that it serves not just as a repository of data, but as a dynamic tool for knowledge creation, best-practice dissemination, and industry connectivity.

4.2.1 Functionality / Services

This section of the thesis explores the functionality and services offered by the energy-bim.com platform, highlighting its innovative approaches to improve user experience and knowledge dissemination in the domain of sustainable construction. The key features and services of the platform are outlined as follows:

Smart Search Facilities: The platform addresses the limitations of current search methodologies, which typically process documents as 'black boxes' and lack effective filtering mechanisms. By implementing advanced search facilities, the platform aims to provide more refined, targeted search results, significantly improving information retrieval efficiency.

Backend Ontology: A fundamental element of the platform is the incorporation of a back-end ontology, which is essential for facilitating smart search functionality. This ontology serves as a structured framework for building and enforcing concepts related to sustainable construction, ensuring that search results are relevant and contextually accurate.

User Profiles: Personalised user profiles are a critical component of the platform, designed to empower users and foster customer loyalty. These profiles facilitate customised experiences, with services such as knowledge and training events that also promote professional networking opportunities.

Bi-Networks for Information Exchange and Enrichment: The platform integrates both electronic and professional networking capabilities, focusing on leveraging both ontology and user profiling. This bi-network system is instrumental in facilitating the exchange and enrichment of information among various stakeholders.

Interface for Shared Resources and Services: The platform offers an interface for accessing a variety of resources and services. This includes both free-to-use resources and premium tools, which may be available on a rental basis. Such an interface is designed to cater to a wide range of user needs, from casual inquiries to more intensive professional requirements.

CPD Facilities: Continuing Professional Development (CPD) facilities form a vital part of the platform's educational resources. These include live and virtual case studies, as well as detailed information on access to professional training programmes, training courses, and e-Learning content. The availability of these CPD resources plays a crucial role in supporting ongoing professional development in the field of sustainable construction.

In general, the energy-bim.com platform stands out for its comprehensive suite of functionalities and services, each meticulously designed to support the evolving needs of professionals in sustainable construction. These features not only facilitate efficient information management and retrieval, but also contribute to the professional growth and networking of its users.

As part of the platform, the researcher has implemented a search service that performs semantic search on the platform BIM knowledge base from a set of authoritative URIs. The submitted BIM query has a set of associated ontological concepts to improve the precision and recall of the returned results. The search service also aggregates data from various trusted sources related to BIM via web crawling. These sources can be proposed by users and validated by experts according to their relevance to BIM for energy efficiency. The list of these sources can be seen below:

1. http://www.bim.psu.edu
2. http://digitalbuilding.lu
3. http://www.list.lu
4. http://objectif-bim.com

5. http://www.batiment-numerique.fr
6. http://www.accept-project.com
7. http://construction21.org
8. http://bimcrunch.com
9. http://mediaconstruct.org
10. http://bimblog.housef
11. http://geometrygym.wordpress.com
12. http://cardiff.ac.uk
13. http://www.ines-solaire.org
14. http://eksergia.fi
15. http://buildingsmart.fi

The researcher has also implemented a Professional Networking Service that enables users to collaborate using social networks such as LinkedIn and Twitter by aggregating associated data. This service also allows users to search for partners and colleagues and identify the corresponding networking profiles based on a set of BIM interests and disciplines.

The *Event Calendar Service* is a reminder of the engineering community's important BIM events. Users can subscribe and synchronise these events relating to sustainability with their personal calendars.

The *BIM Tools Service* was implemented to expose a number of BIM tools addressing various aspects of energy, such as carbon emissions, energy simulation, etc.

The *BIM Training Service* was implemented to enable users to identify courses and lectures related to BIM for energy efficiency in construction from various institutions such as universities, research organisations, government agencies, etc.

Within the energy-bim.com platform, the query and expand methods are exposed to provide the key use of the ontology to drive the search engine. First, the terms within the ontology provide keyword suggestions when entering search terms (using the query method). Secondly, the relationships between terms help users expand/restrict their queries based on suggestions from the ontology (as observed in Fig. 13).

The innovative dimension of the energy-bim.com platform lies in its open, scalable, and polymorphic context-based widgets that reconfigure and update themselves to respond to changing user context and (BIM-related) queries while enabling serendipitous BIM information and knowledge discovery. Each service has a corresponding widget that users can update and administer remotely.

Analysing user statistics and comments helped identify some key issues to be addressed in future

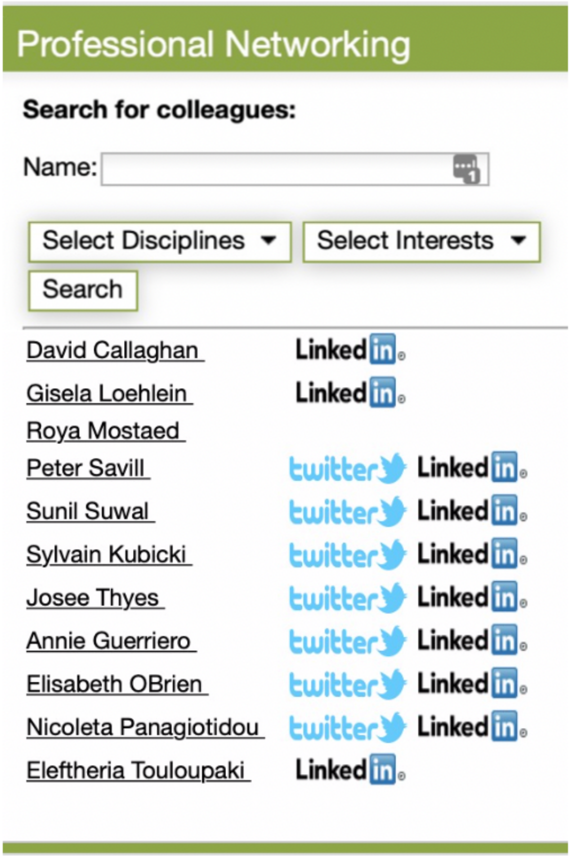


Figure 12: The professional network service

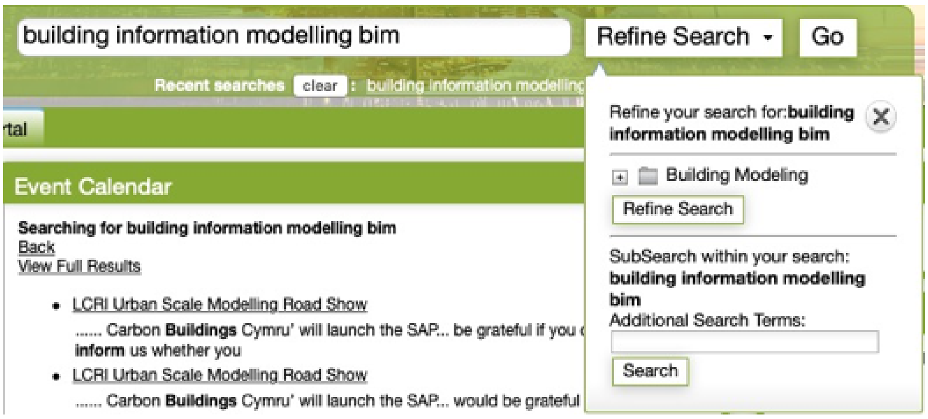


Figure 13: The energy-bim platform search system

platform releases.

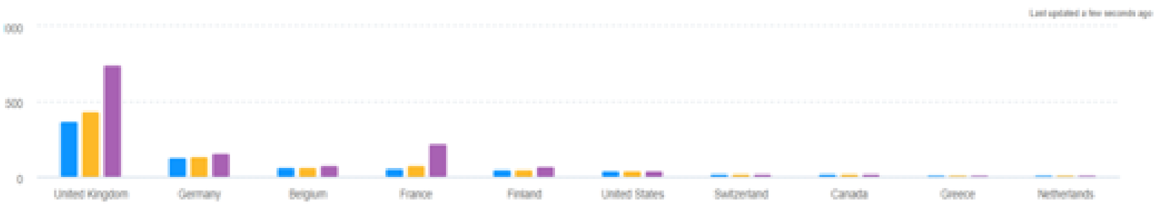
The researcher has exposed the energy-bim.com platform as an online location to create a community of users in the field on BIM training for energy efficiency. From the monitoring interval between December 2017 and February 2018, the researcher has attracted new users and identified increased visitors. Using the Woopra analytics illustrated in Fig. 14 above, the researcher collected several statistics on the web activity of the platform. Fig. 14 illustrates the total number of visits to the energy-bim.com platform over a trial period of 3 months. It can be identified that popular services and the associated accessed content are related to different European countries.

Regarding the online activity of accessing BIM resources, Fig. 4 illustrates that a significant percentage of users have returned to conduct web searching for BIM resources. Analysing the geographical provenance of visitors, the researcher has identified that the platform is of interest not only to UK visitors but also to the US or other EU countries.

Initial statistics indicate that the proposed web portal:

1. has the potential to engage further practitioners in the delivery of BIM interventions as provided by the work on portal validation;
2. it contributes to the ongoing discussion and integration of BIM in energy efficiency.

In addition, technological diversity can contribute to the emergence of new business models and help develop the online marketplace for the construction industry. The researcher developed an



Countries	People	Visits	Actions
United Kingdom	372	439	740
Germany	131	134	160
Belgium	61	64	81
France	57	79	224
Finland	47	48	70
United States	44	44	44
Switzerland	22	22	24
Canada	20	20	20
Greece	17	17	17

Figure 14: The visitors of the energy-bim.com platform

effective and easy-to-use user interface for a BIM training platform for the energy sector in order to improve the *energy-bim.com* platform to refine the search and display of results and the online support and interface features.

Based on the critical review of the related literature, there is a clear gap in which no study has combined the analysis of secondary sources of evidence with social media information, using TF-IDF, association, correlation and cluster analysis to infer roles and responsibilities regarding the adoption of BIM in the industry.

4.2.2 Harvesting BIM data from web sources

In the form of interactive, responsive and user-orientated applications utilising the latest technologies, the *energy-bim.com* portal provides access to integrated BIM resources for a community of users and professionals. The framework is an open, scalable, polymorphic context-based solution with modules that allow BIM information and knowledge to be unlocked through a symbiosis of technology [284].

Fig. 8 presents an overview of the *energy-bim.com* portal by highlighting its primary widgets (Building Information Modelling, Sustainable Sourcing, Training, Sustainable Design Tools, Web Search, and Professional Networking) in conjunction with the advanced search facility that filtrates the search results of queries directed to the back-end ontology.

The platform retrieves and stores a repository of BIM data sources, including:

1. Documents such as scientific publications, standards, and regulations in the field of BIM for energy efficiency;

2. Twitter data from professional organisations, bodies, and actors working in the field of BIM and energy efficiency

The researcher has retrieved around 80 key BIM-related publications and regulations alongside 4 million tweets for input within the analysis. The researcher has developed a search service that searches the BIM knowledge base from several authoritative URIs (Uniform Resource Identifiers) as part of the application. The BIM query submitted to the publication search engines contains several related ontological principles to improve accuracy and retrieve the returned data. The search service also includes data from reputable sources related to BIM via the Web. These sources can be suggested by users and tested in terms of energy efficiency by a group of experts.

The scalable, multi-focal, context-based widgets of the *energy-bim.com* platform are the creative aspect which can be reconfigured to serve the following purposes:

1. answering evolving user contexts and (BIM-related) questions

2. allowing for smart BIMxt (Building Information Modelling Extension Network) information and the exploration of knowledge

Every service has a related toolbar that users can update and control remotely. Analysing usage data and feedback have helped to define several key problems to address in future product launches.

4.2.3 Objectives

4.2.3.1 Frequency determination

Our first objective was to determine the frequency of each BIM role in the corpus of 60,000 tweets and sentences in the 80 scientific publications. In contrast, the second objective was to determine the frequency of each BIM skill in the filtered corpus of data. The frequency of a role or skill has been determined based on the formula below:

$$TF(t) = \frac{frequency\ of\ t}{total\ terms\ of\ the\ same\ type\ as\ t} \tag{1}$$

From Equation 1, t is a skill or a role, and $type\ of\ t$ can be a skill or a role. This has facilitated the classification of skills and roles in order to understand their importance in the overall dataset.

4.2.3.2 Association determination

Our third objective was to determine the associations between the BIM roles and skills with their associated importance.

```
userList = SELECT DISTINCT FROM users
INITIALISE combinationList, roleSkillCombinationMatrix, skillRoleCombinationMatrix
for each user in userList do
    INITIALISE currentUserRoleList
    for each role, skill do
        ADD role TO currentUserRoleList
    end for
    INITIALISE currentUserSkillList
    for each skill do
        ADD skill TO currentUserSkillList
    end for
end for
for each combination in combinationList do
    for each role, skill do
        if role, skills IN combination then
            ADD role, skill TO roleSkillCombinationCountMatrix
            roleSkillCombinationCountMatrix[role][skill from combination] += 1
        end if
    end for
end for
```

Figure 15 Roles and skills correlation + combination matrix algorithm

Fig. 15 illustrates the algorithm the researcher used to determine the roles and skills associated with the data set. We took the raw database and the list of previously determined roles and skills as input. The researcher then created two combination matrices. Each element in *roleSkillCombinationCountMatrix* has been used in Equation 2, and each element in *skillRoleCombinationCountMatrix* has been used in Equation 3. In the context of scientific publications, we considered the role and skill to be associated if they could be found in the same sentence.

4.2.3.3 Importance determination

Our fourth objective was to determine the importance of each role and skill, taking into account the average number of occurrences in conjunction with all its counterparts as in Equation 2.

$$Importance\,(skill) = \frac{1}{n}\sum_{n=1}^{n}(no.\ of\ occurrences(association(skill + role)) \tag{2}$$

, where n is the number of distinct associations between a particular skill and multiple different roles.

$$Importance\,(role) = \frac{1}{n}\sum_{n=1}^{n}(no.\ of\ occurrences(association(role + skill)) \tag{3}$$

, where n is the number of distinct associations between a particular role and multiple different skills.

4.2.3.4 Correlation determination

Our final objective was to determine the degree of correlation between the terms (roles and skills) occurring within the context of the same tweet or sentence. The researcher achieved this by determining the semantic relationships between concepts applying the metric cluster method [28], which factors in the number of co-occurrences of concepts with their proximity in the text. To determine ontological dependencies between concepts, the researcher has applied a combination of TF-IDF and Metric Clusters methods to infer the correlation factor between BIM roles and skills [301].

$$C_{\mathrm{u,v}} = \sum_{k_{\mathrm{i}}\in V(S_{\mathrm{u}})}\sum_{k_{\mathrm{j}}\in V(S_{\mathrm{v}})}\frac{1}{r\left(k_{\mathrm{i}},k_{\mathrm{j}}\right)} \tag{4}$$

$$C_{\mathrm{u,v}} = \frac{1}{\mathrm{Min}\left[r\left(k_{\mathrm{u}},k_{\mathrm{v}}\right)\right]} \tag{5}$$

The correlation technique is presented in Equation 4, where the distance $r(k_i, k_j)$ between two keywords k_i and k_j is calculated as the number of words interjected between two terms in the same tweet or sentence. $V(S_u)$ and $V(S_v)$ represent the tuples of keywords with the stems S_u and S_v associated with them. To improve the degree of generalisation, the application instances of this formula in the experiments discarded different semantic discrepancies of concepts within the Twitter text. Instead, they used a more efficient correlation technique as in Equation 5, where $r(k_u, k_v)$ represents the minimum number of words that separate the two keywords k_u and k_v in each tweet (or sentence).

The hypothesis when applying Equation 4 was that a low difference between the correlation factor and the value of 1 for the denominator represents a high correlation strength between the two terms.

In order to establish a correlation factor threshold, the researcher decided to limit the application of the formula to instances of tweets that embed both terms. The researcher specified the following exception: *Should the minimum distance between two analysed keywords be null, the default correlation factor will still be "1".* In this exception, the correlated terms are considered candidates to be part of a new unified term.

The development and refinement of the BIM training platform involved a systematic approach consisting of gathering requirements, technical implementation, and rigorous testing. This section provides an in-depth look at this process, outlining how the platform evolved iteratively to address user needs through robust software engineering and validation practices.

First, the requirements were gathered based on exhaustive user research, including surveys and interviews with construction professionals. This provided critical insights into their pain points, challenges, and goals for a BIM training solution. The requirements spanned functionality, content, delivery formats, collaboration features, assessment capabilities, integration with industry tools, and more.

Next, the platform was developed through an agile microservices architecture using a cutting-edge technology stack. Each component was built to be modular and scalable using technologies like Node.js, React, MongoDB, and Docker. Custom features were engineered to enable interactive 3D visualizations, offline mobile access, secured authentication, and more.

4.3 Requirements, Technical process and testing

Extensive testing activities were conducted in parallel throughout the development lifecycle: automated unit, integration and end-to-end tests validated functionality. Manual usability testing provided qualitative user feedback. A/B tests compared design variations. Canary releases detected defects. Security testing identified vulnerabilities.

This combination of understanding user requirements, purpose-built technology architecture, and rigorous validation was instrumental in creating an engaging platform tailored for construction learning. The approach ensured that industry needs were fully addressed through well-engineered and thoroughly tested software capabilities. The subsequent sections further detail each aspect of this user-driven development and testing methodology.

To meet the needs of diverse construction industry learners, the following key user requirements were identified:

Requirements Search for BIM resources filtered by role, skillset, discipline and application area, such as energy efficiency or sustainability. This caters to different learners' interests and responsibilities. Training content tailored to different proficiency levels ranging from beginners to advanced experts. This provides appropriate challenges and progression. Delivery via text, images, 3D models, video tutorials and webinars to accommodate different learning styles. Online forums for peer collaboration enable knowledge exchange between professionals across organizations.

- Assessments to evaluate progress with AI-driven recommendations for personalized learning pathways aligned to development goals.
- Dashboards to visualize skill profiles, track competency development, and receive suggested training assignments.
- Integration with industry BIM software tools and open standards for practical applied learning.
- Responsive mobile-friendly interface and offline access features to support on-site learning.
- These requirements aim to provide an inclusive platform for diverse construction education needs while aligning with industry workflows. The ultimate goal is scaled and accelerated adoption of BIM methodologies across the workforce through engaging and targeted learning experiences.
- Secure access control and user identity management to protect proprietary organizational resources.

Technical Process The BIMEET platform backend was developed using a Node.js runtime environment for scalability, with a MongoDB NoSQL database chosen for flexible and dynamic storage of unstructured BIM content. Elasticsearch search engine was integrated to enable robust and efficient querying across structured and unstructured data. The frontend user interface was built using the React framework for smooth responsiveness across devices and Redux for centralized state management. Material UI React component library provided accessible and consistent UI elements compliant with industry standards. Custom SVG illustrations and THREE.js were used for interactive 3D visualizations of BIM models tailored for web delivery.

Video content was adapted for web streaming using HLS (HTTP et al.) adaptive bitrate protocol to adjust quality based on network conditions. Offline access features leveraged service workers and IndexedDB browser storage for reliability in low connectivity.

Security was implemented using JSON Web Tokens and role-based access control, ensuring protected access to proprietary BIM resources. Docker containers provide standardized packaging for portability across environments.

A continuous integration and deployment pipeline was set up using GitHub Actions workflows to automate testing and delivery of updates. Unit testing was performed with Jest and React Testing

Library. End-to-end testing involved a Selenium browser automation suite replicating user workflows. The microservices architecture allowed individual platform capabilities to be developed, deployed and scaled independently. Core services included user management, content management, search, analytics, discussion forums, assessment delivery and learning analytics.

This combination of technologies enabled an engaging, responsive and secure platform tailored for construction learning. The iterative approach with continuous testing and DeliveryFacelift feedback facilitated the platform's evolution to meet emerging industry needs.

Testing Several manual testing techniques were utilized throughout the development process to validate functionality, usability, and robustness. User interface testing was conducted manually by replicating all possible user workflows to uncover design flaws or bugs. Exploratory testing techniques such as error guessing and boundary analysis were used to identify corner case defects. Focused user testing sessions involved representative users from the construction industry performing realistic tasks and providing feedback. A usability heuristics assessment was performed to evaluate the platform against established usability principles. Integration testing was manually conducted between related components. End-to-end workflow tests validated the platform's support of comprehensive real-world use cases. Security testing techniques such as authentication bypass and input fuzzing were manually performed. This diverse set of manual testing approaches identified defects early, and the platform was refined to provide an intuitive and engaging user experience specifically tailored to construction industry learners.

4.4 Data collection and filtering

The researcher uses a generic mining function $f(t)$ $f(t) : C_t \rightarrow \{R, S\}$, where $C_t = \{t_{e1}, t_{e2}, ...\ t_{em}\}$ is a collection of endpoints (Twitter accounts), and each t_{ei} is a Twitter account which generates a set of tweets, $R = \{r_1, r_2, ...\ r_b\}$ is a set of retrieved roles, and $S = \{s_1, s_2, ...\ s_c\}$ is a set of skills recovered.

During the collection process, each tweet must be pre-processed to make it readable for the parsing algorithms and regular expressions. In a scientific publication, a sentence with BIM-related keywords is the equivalent of a tweet. Furthermore, this process increases the reliability of the results and the reduction of bias (for example, a role is present as part of a *camelCase* construction).

Table 3 presents the portfolio of companies and organisations that have been used as primary data sources for the extraction of social media.

The capturing process for the tweets and scientific publications was supported by a list of organisations identified from three sources:

1. IP detection and organisational identification forensic algorithms;

2. Followers of the "@BIMEET" Twitter account;

Name of organisation	**Twitter account**
Group CSI	https://twitter.com/groupecesi
INESs Solaires	https://twitter.com/ines solaire
BRE Academy	https://twitter.com/BREAcademy
Écoles des Ponts Paris Tech	https://twitter.com/EcoledesPonts
ESTP	https://twitter.com/estpparis
Universite de Lièege	https://twitter.com/UniversiteLiege
UC Louvain	https://twitter.com/UCLouvain be
Citt'a di Modena	https://twitter.com/cittadimodena
ORSYS Formation	https://twitter.com/ORSYS
BEC partners SA	https://twitter.com/becpartners
Middlesex University	https://twitter.com/MiddlesexUni
House of Training	https://twitter.com/Houseoftraining
Sapienza Universita	https://twitter.com/sapienzaRoma
Scuola Master F.lli Pesenti, Politecnico di Milano	https://twitter.com/master pesenti
Le Moniteur	https://twitter.com/Le Moniteur
Technical University of Denmark	https://twitter.com/DTUtweet
Norwegian University of Science and Technology	https://twitter.com/ntnu
UIC Barcelona	https://twitter.com/UICbarcelona
Mensch und Maschine	https://twitter.com/MuMDACH
Zigurat	https://twitter.com/Ziguratdigital
BIMEET EU	https://twitter.com/bimeetEU
H2020EE	https://twitter.com/H2020EE
H2020 BIM plement	https://twitter.com/H2020BIMplement
ECTP Secretariat	https://twitter.com/ECTPSecretariat

Table 3: List of companies/organisations used as primary data sources and their Twitter account

3. BIM institutions indicated by partners.

The researcher considered these authoritative organisations highly relevant as driving forces of the continuous transformation of the BIM industry as they are active users of the *energy-bim.com* platform. The researcher has used the 60,000 filtered tweets from the total corpus of 4 million posted by the portfolio of companies presented in Table 3.

For the filtering process, the researcher uses a set of expressions, $N_e = \{exp_1, exp_2, ...exp_n\}$, where exp_i is an expression defined in the list of expressions presented below, filtering $C_t^{''}$, $C_t^{''} \subset C_t$, where custom tuples of roles and skills are defined as $C_t^{''} = \{g(t_{eij}), \rightarrow \{r_i, s_j\} \in \mathbb{N}_E\}$

The following set of regular expressions (RegExp) and pattern matching rules was needed to filter the data due to the noisy nature of social networks, in which casual spelling, grammar, and brief expressions are often used. The same expressions were also used to filter data found in scientific publications, a process further aided by manual recognition.

```
+((contractor\manager\designer\engineer\client\)|\skills|\).+ (\energy| \construction)
+((\BIM\construction\energy)|\skills|\).+ (\need| \require)
+((\BIM\construction\energy)|\roles|\).+ (\need| \require)
+((\BIM\construction\energy)|\actors|\).+ (\skills| \competencies)
+((\BIM\construction\energy)|\knowledge|\).+ (\requirements| \require)
+((\BIM\construction\energy)|\skills|\).+ (\need| \require)
+((\BIM\construction\energy)|\competencies|\).+ (\need| \require)
+((\skills\competencies\knowledge\expertise)|\BIM|\).+ (\energy| \construction)
```

As representative starting points, the concepts of "skills" and "roles" were created for regular expressions with the primary objective of filtering the inaccurate context before analysis. The roles were frequently associated with keywords that included "construction", "skills", and "energy". Several terms, including "training" and "knowledge", could be found in sentences that encompass both roles and skills.

4.5 Establishing association, importance, and correlation

The following functions are also used to determine the associations, correlations, and importance between roles and skills:

a) $f(a)$ for determining the association between the roles and skills facilitated by TF-IDF techniques (as described in Equation 1): $f(a) : [R, S] \rightarrow [RS_a^{'}, A]$, where $RS_a^{'}$ is a set of pairs of roles and skills with a Boolean value A that indicates whether the pair of roles and skills is associated or not.

b) $f(i)$ for determining the importance (as described in Equations 2 and 3): $f(i) : [R, S] \rightarrow [RS_c^{'}, I]$, where $RS_i^{'}$ is a set of pairs of roles and skills with their importance value I.

c) $f(c)$ for determining the correlation through the proximity concept (as described in Equations 4 and 5): $f(c) : [R, S] \rightarrow [RS_c^{'}, F]$, where $RS_c^{'}$ is a set of pairs of roles and skills with their correlation value F.

4.6 BIM roles frequency

In this experiment, the researcher seeks to analyse roles based on their frequency in the document corpus. Figure 16 presents a classification of roles based on their frequency, where roles such as "energy modeler" and "energy procurement specialist" have been identified as important based on the analysis. A lower frequency has been attached with roles such as "HR specialist", "finance specialist", and "communication officer", demonstrating a key area of interest for developing more BIM competencies and educational programmes.

From Figure 16, it can also be observed that BIM has an impact on different areas of practice ranging from "clerical" roles to "scientists", which brings a relatively high diversity of BIM roles.

It can be observed that the repository of scientific publications incorporates a greater variety of roles than the social media corpus due to the presence of lower-skilled roles such as "plasterer", "dry liner", and "bricklayer". However, as expected from the difference in the size of the repositories, the frequency of the roles found in scientific publications is lower. However, we can observe the mutual presence of several high-skilled roles such as "BIM teacher", "facility manager", "building professional", and "quality assurance manager".

From the construction industry perspective, this experiment demonstrates that disciplines and roles need to adapt to emerging challenges identified in the industry, with a key emphasis on the roles and competencies related to "energy" that present the highest frequency in the analysis.

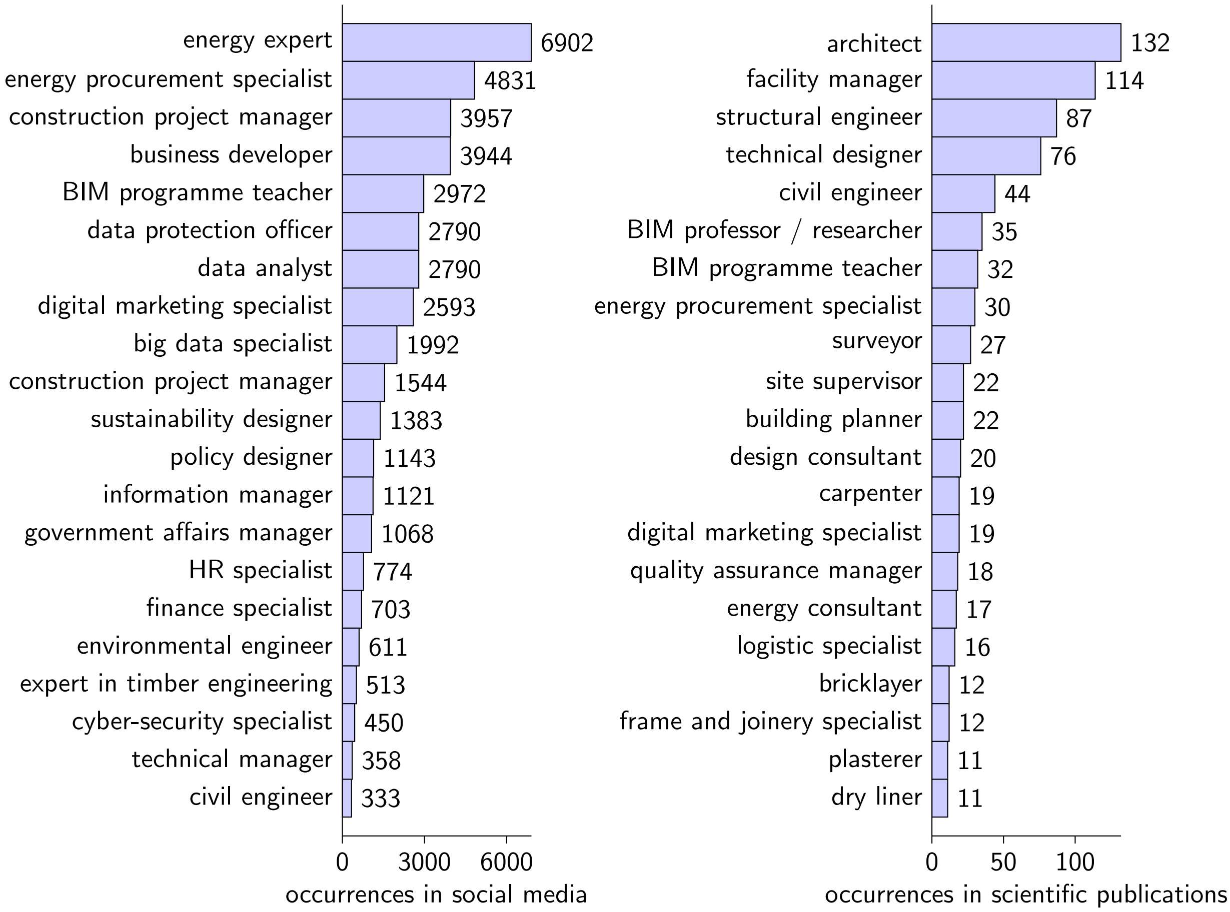


Figure 16: Frequency of BIM roles on social media (left) vs. scientific publications (right)

4.7 BIM skills frequency

In this experiment, the researcher identifies skills in relation to their frequency. Figure 17 presents a classification of skills based on their frequency, where skills such as "4D simulation", "construction", and "urban planning" have been identified as important from reports in user tweets. A lower frequency has been associated with skills such as "structural engineering software" and "structural analysis".

From Figure 17, it can be identified that BIM has a direct competency relation to advanced skills such as "urban planning" and "4D simulation".

The mutual presence of specific skills can be observed in both repositories, such as "Revit operation", "coordination skills" and "energy management", while also noticing a more extensive variety of skills in scientific publications, ranging from management to highly specialised roles such as "solar and wind panel operation". As expected, due to the difference in repository size, we can notice a lower frequency of roles from scientific publications.

It can be seen that the skills with the highest frequency are "4D simulation", "construction", and "urban planning". These are followed by "training skills", "leadership", and "management", all of which are highly frequent. Skills directly correlated with the "energy" work field are in the middle range of the spectrum. It can also be observed that skills related to *business & management* such as "marketing skills", "speaking skills", "entrepreneurial skills", and "cooperation skills". The results also report new skills related to *technology* such as "ICT", "IoT", "virtual reality operation", "Revit operation", "automation" and "analytics".

The experimental results also report the need to advance the development of new skills such as "4D simulation" as well as a greater understanding of and Communication Technologies (ICT) and IoT advancements, along with their applicability for the Construction industry. These findings show an intrinsic transformation of the construction industry and the advantages of digitisation for supporting more informed and efficient energy practices.

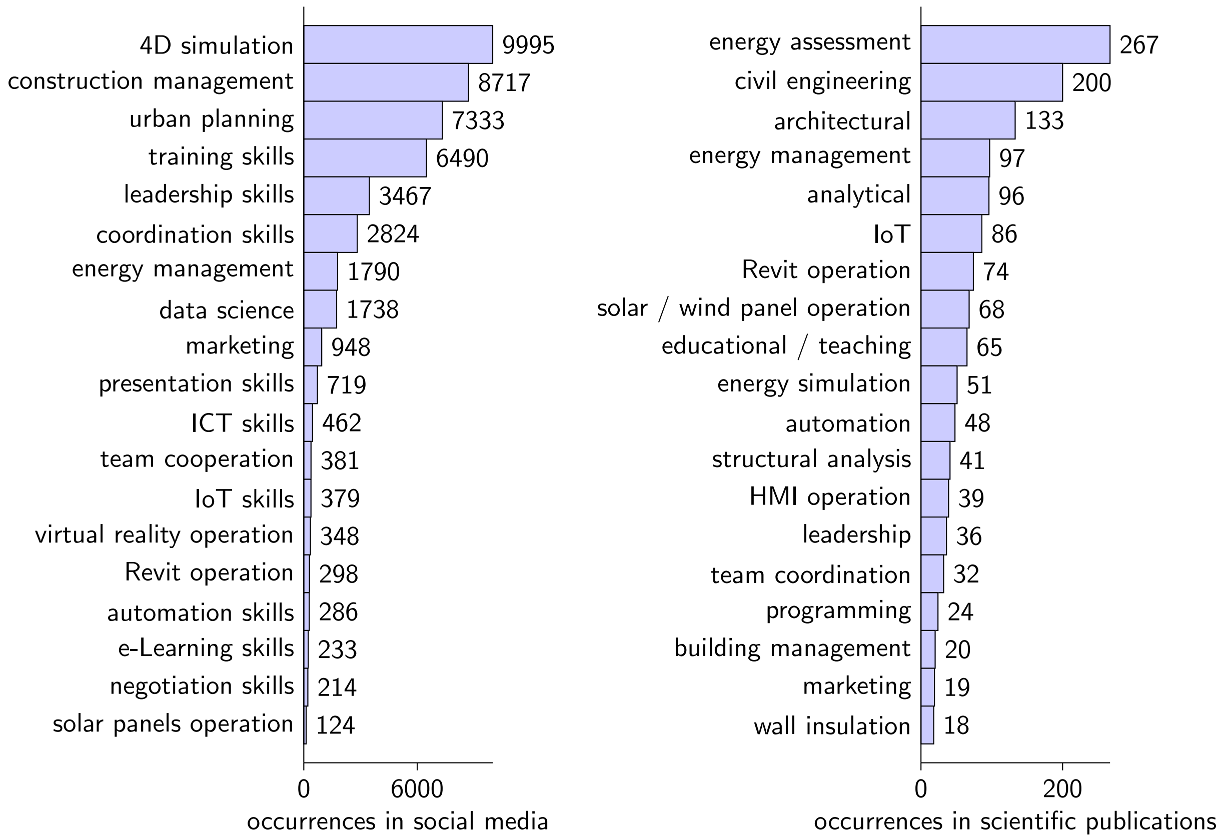


Figure 17: Frequency of BIM skills on social media (left) vs. scientific publications (right)

4.8 Importance of roles in relation to skills by category

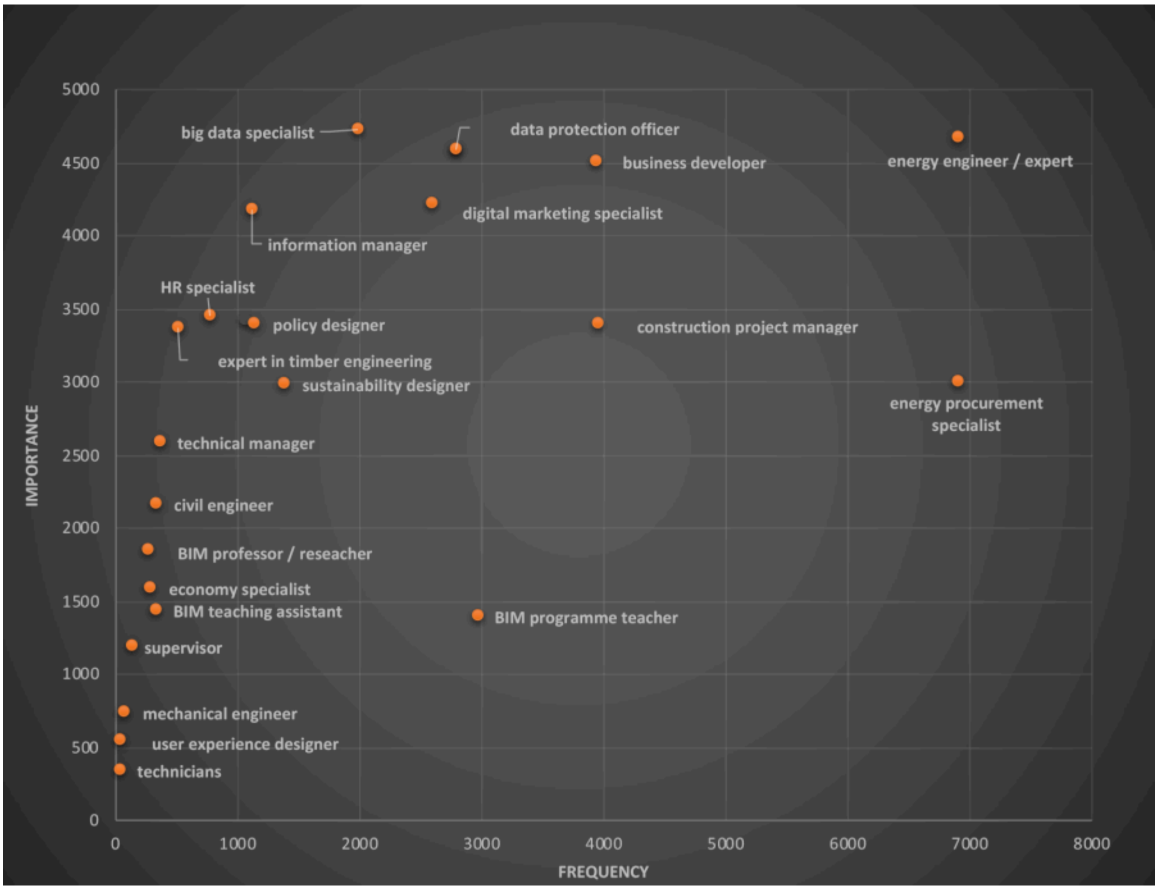


Figure 18: Frequency and importance of roles in relation to skills

The main objective of this experiment was to determine a correlation between the importance and frequency of each role and skill. Importance and frequency are defined by the formulas presented in the previous subsection. In Figure 18, the researcher has provided the results of the correlation analysis between frequency and importance of roles and skills, with frequency on the X axis and importance on the Y axis. In general, a positive relationship between frequency and importance can be observed.

For energy-specific roles, it can be observed that the role that has both the highest frequency and the highest importance is "energy engineering/expert", followed by "energy procurement specialist", while "information manager" has high importance but low frequency, followed from here by "economy specialist".

Regarding management roles, the roles that have the highest frequency and importance are "digital marketing specialist" and "construction project manager". The role of high frequency but low importance is "architecture & construction project manager". For research roles, the most important roles are "BIM professor/researcher", "BIM teaching assistant" and "BIM programme teacher".

In relation to technology roles, the high-importance roles are "data protection officer", "digital marketing specialist", and "big data specialist" with lower frequency for "user experience designer" and "technicians".

4.9 Frequency and importance of skills in relation to roles

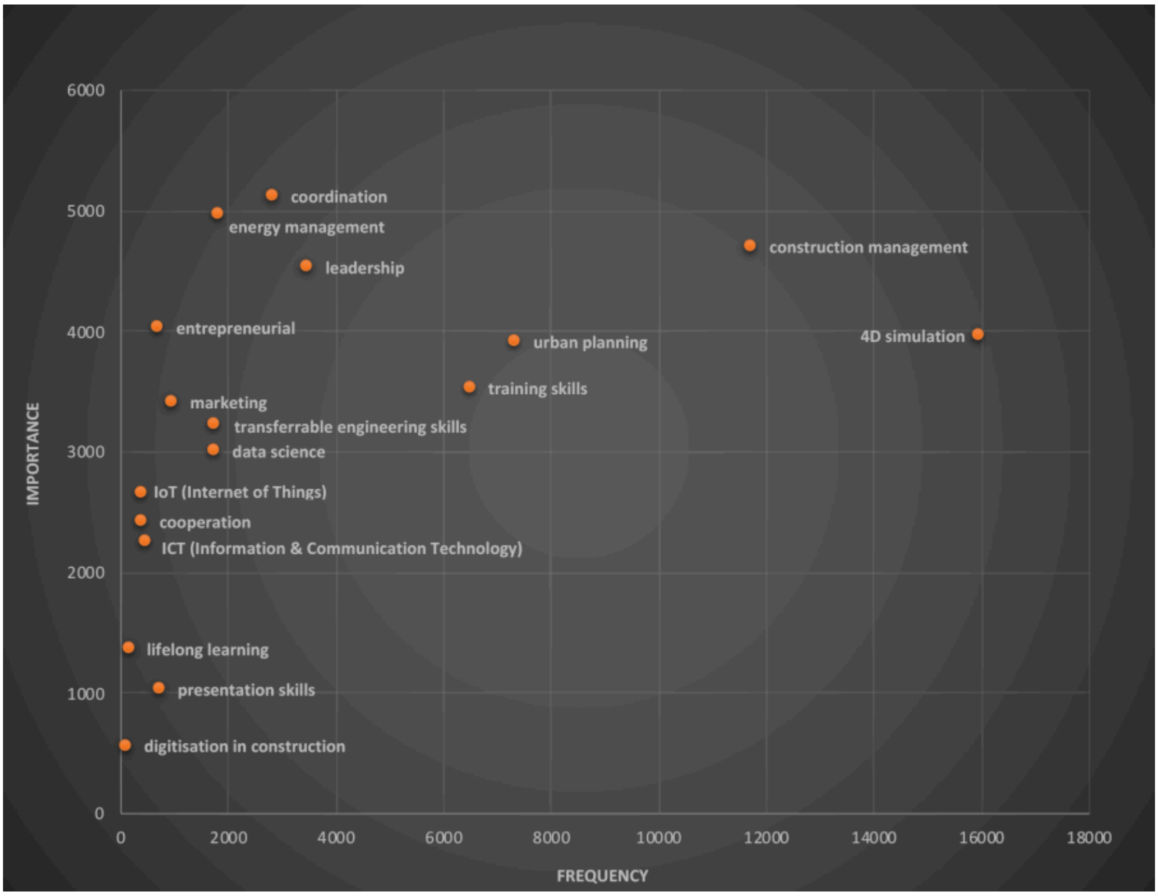


Figure 19: Frequency and importance of skills in relation to roles

The resulting analysis depicts the correlation between frequency and importance in relation to

BIM roles and skills, with frequency on the X axis and importance on the Y axis. From the results, a linear trend can be observed with respect to the correlation between frequency and importance. Skills such as “coordination”, “energy management”, and “leadership” are characterised by high importance but relatively low frequency. At the midpoint of the importance and frequency scale, skills such as “urban planning” and “training skills” can be observed, along with skills with similar low frequency, such as “IoT”, “cooperation” and “ICT”. At the lowest end of the scale, very low frequency skills can be observed, such as “lifelong learning”, “presentation skills”, and “digitisation skills”.

The results report numerous trends in relation to the expertise and competencies required in the construction industry. This expertise can be obtained by developing a set of skills and roles around BIM and energy, supported by big data technologies to tackle digitisation and improve sustainable practices of the industry.

4.10 Correlation factor between roles and skills

In addition, the researcher has determined the correlation factor between the top 100 combinations of roles and skills and converted the results into the format of an ontology, having the roles and skills as classes and the correlation factor as the property of the object.

In Figure 20, the presence of some central roles and skills with a large number of edges, such as the “energy storage committee”, “construction”, “management”, “energy transition”, “energy performance”, and “energy management” can be observed, along with a group formed around the nodes containing the “energy” keyword, with “energy storage committee” being the most transited node.

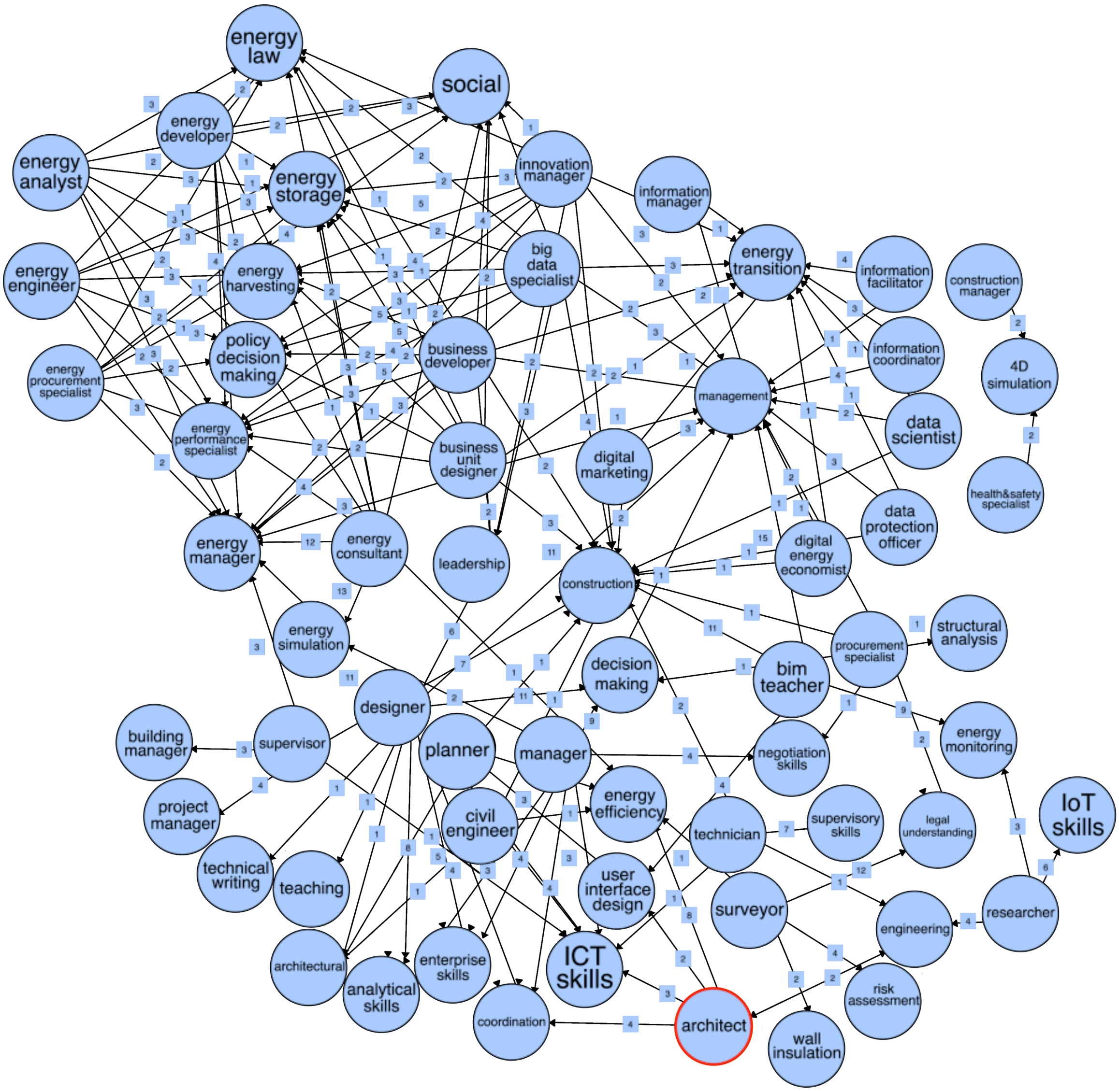


Figure 20: Ontology visualiser for correlation between BIM roles and skills

The researcher has also calculated a correlation factor according to Equation 5. It can be seen that many roles and skills have a maximum correlation factor of 1 (100%). These correlations illustrate a number of industry migration and transformation trends that identify the areas of the construction industry where new training and education programmes are required.

Semantic analysis provides a clear illustration of the key concepts involved in the promotion of BIM for energy efficiency and their dependencies based on a correlation factor metric. This type of analysis can facilitate a holistic representation of possible areas of improvement that need to be addressed in the construction sector. Through the use of ontological conceptualisation, the researcher has identified existing informational clusters that reflect various levels of digitalisation in the sector.

This chapter sets the objective to address the first research question: *How could an existing web application be enhanced using context to maximise User Experience?* This is achieved by implementing a novel text mining approach that analyses social media in addition to secondary sources of evidence to establish correlations between BIM roles and skills. Using ontological dependency analysis to inform training and educational programmes has helped to understand the degree of correlation of skills with roles. The research produced a web-based semantic mining environment that determines BIM roles and skills, as well as their correlation factor, with an application for energy efficiency. The researcher also shows that (a) construction skills and roles are highly dynamic and evolve continuously, reflecting the construction industry's digital transformation, and (b) the importance of socio-organisational aspects in construction skills and related training provision.

5 Delivering context-informed Digital Twins through semantics

5.1 Introducing the CUSP platform

This chapter seeks to elucidate the second research question of the thesis: *How to derive context from static artefacts (buildings) and dynamic artefacts using a Digital Twin?*

In the pursuit of answering this, the attention shifts towards the semantic structures underpinning the Computational Urban Sustainability Platform (CUSP).

At the commencement of the "Delivering context-informed Digital Twins through semantics" section, it is pivotal to acknowledge the centrality of user involvement that underpins the participatory approach of this research. From the outset, the development process has been deeply collaborative, engaging users directly to shape the trajectory of the bespoke UX model and the digital twin platform. This participatory ethos underscores the project's commitment to ensuring that the solutions devised are not merely technical showcases but are intrinsically designed with the user's experience and expertise at the forefront. It is the insights, feedback, and active contributions from the users that have been instrumental in steering the development towards a platform that is not only technologically sophisticated but also intuitively aligned with the real-world needs and contexts of the end-users. This user-centric approach exemplifies the research's core principle: that the most effective technological innovations are those co-created with their intended beneficiaries, guaranteeing that the tools we develop are as accessible and relevant as they are advanced.x5

The software architecture of this system comprises three core components: *IoT sensors*, *semantic middleware*, and *User Interface services*. IoT sensors installed within the built environment capture data on various parameters related to energy. The semantic middleware, equipped with the OWL ontology, interprets these raw data in a semantically rich context-aware format. This middleware deciphers sensor data and identifies interrelationships and dependencies between parameters, thus providing a comprehensive view of energy performance.

The processed data from the middleware is then made available to end-user services, which include a diverse range of applications like energy monitoring dashboards, operational control systems, and energy optimisation tools. This architecture also allows for potential integration with machine learning algorithms, leveraging the semantically rich data for predictive analytics and optimisation models.

The user experience (UX) of the system focuses on ease of understanding and actionable insights. The ontology-based interpretation of sensor data simplifies complex information, allowing users to quickly comprehend energy performance and identify potential areas for improvement. The User Interface visually represents these insights, highlighting key metrics and trends. Together, these features ensure an intuitive and engaging user experience that enables users to make informed decisions about energy use and conservation. Future improvements aim to further enhance the UX design based on user feedback and evolving energy management needs.

The author has not developed the semantic middleware, and hence, has not elaborated within this thesis. On the contrary, the author has played a crucial role in advancing the smart city and energy performance management solution by actively contributing to the design and implementation of the Web User Interface (UI). This interface is a critical component to enable user interaction with the system and visualisation of complex data sets related to urban energy management. Furthermore, the author has successfully integrated the UI with the underlying middleware layer by utilising the Representational State Transfer (REST) Application Programming Interface (API). This integration ensures efficient communication between various architectural components, ultimately leading to an effective and user-friendly solution for monitoring and optimising energy performance within the smart city context.

Semantic frameworks developed within the Computational Urban Sustainability Platform (CUSP) exhibit significant potential for efficient conversion into visually annotated data structures embedded within a web interface. These frameworks represent an innovative step in data presentation, offering a multifaceted approach to data visualisation that surpasses traditional methods.

The CUSP platform represents a novel approach to designing and implementing complex systems modelling. It was born from the recognition that urban systems, with their diverse components and intricate interconnections, require a flexible, scalable, and semantically rich framework for representation and analysis. In this chapter, we delve into the engineering philosophy that underpins CUSP and explore the technical rationale behind the architecture and its associated technologies.

This chapter explores how the semantic structures of CUSP and the associated annotated visual data contribute to the derivation of context from static and dynamic artefacts. In doing so, it provides key insights into enhancing the platform's intuitiveness and seamless user experience, ultimately bolstering the potential of CUSP as an integral component within the smart city's energy management landscape.

Access Control List (ACL)-based security is widely supported in relational databases such as

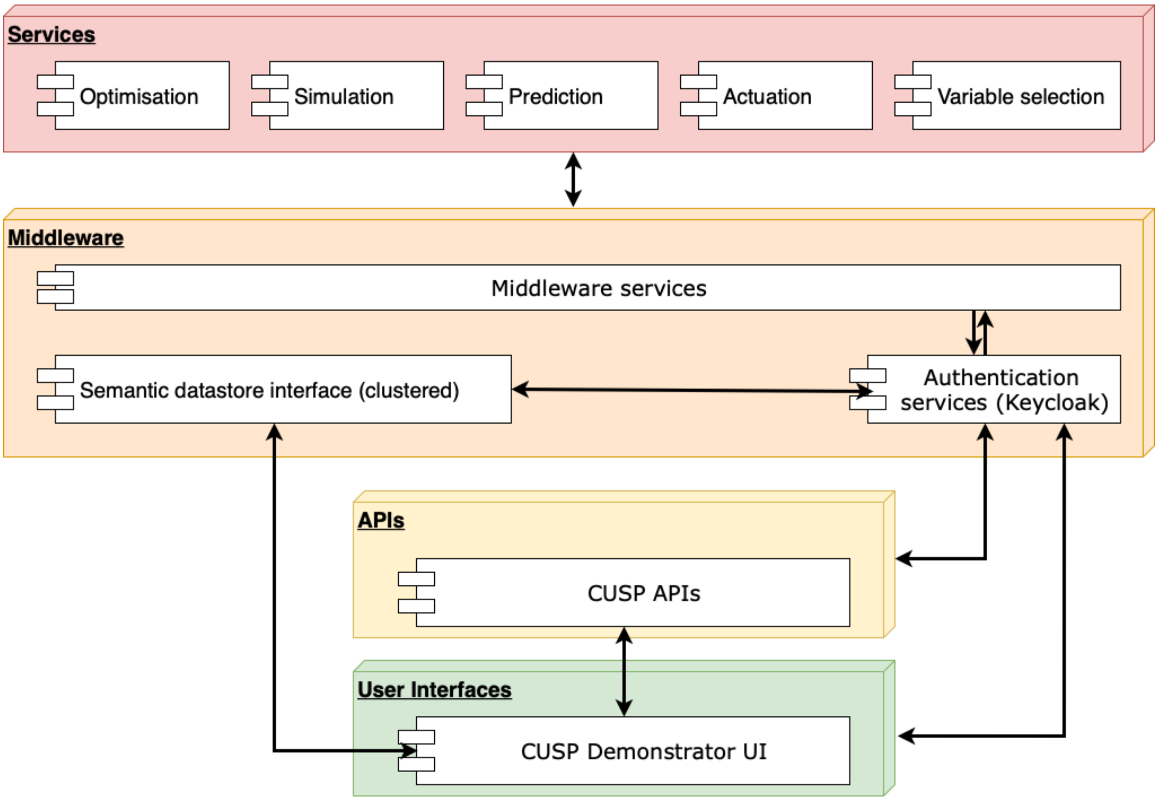
MySQL. Neo4j does not come with pre-installed security features, as even though Access Control List (ACL) security mechanisms such as user support are built into the Neo4j website, maintenance of these mechanisms is left to the application. For the reasons previously mentioned, the researcher decided to rule out the possibility of implementing the Computational Urban Sustainability Platform (CUSP) platform backend on the backbone of any Graph Query Language.

Cities and districts are at the forefront of sustainability objectives. This recurring theme is increasingly being addressed by improving operational efficiency through smart cities and intelligent management solutions adapted to the technical levels of expertise of stakeholders and policymakers. Several researchers are developing the Computational Urban Sustainability Platform (CUSP) at Cardiff University as an immersive decision support tool built to deliver robust urban analysis, aiming to meet this challenge. CUSP enables interactive monitoring and informed decision making through an intuitive user interface. CUSP forms the basis for the third generation of web portals: the semantic Web of Things.

The core of CUSP is a semantic middleware powered by AI and optimisation. It is a vital asset, as it can manage time series data while providing value-added services to maintain the consistency of the Digital Twin. CUSP is also a Digital Twin platform that can incorporate various AI, optimisation and simulation tools, which can be automated, limiting the user participation in keeping the platform and the actual Digital Twin up running. The middleware can manage large volumes of real-time or time series-based data through semantic models such as ontologies, including BIM.

As with any product characterised by high complexity, multiple development iterations, followed by multiple evaluation and benchmarking sessions, were carried out before achieving a stable and scalable version of CUSP. Semantics was incorporated into the code in the pioneering versions of the software. This approach was evaluated as unsuitable in the long run, mainly due to issues with the scalability, readability, and collaboration facility.

5.1.1 Drawbacks/caveats of a semantics-only model



Despite their numerous advantages and the praise of industry leaders, open-source data formats used by semantic models such as OWL have several drawbacks. One of the most prominent examples is that OWL cannot be used to directly access the data store or databases, resulting in unnecessary technical costs and reading/writing delays. The researcher mitigated performance bottlenecks throughout the CUSP platform development process by gradually converting and migrating OWL-based resources to a graphical database.

As previously stated, only some semantics embedded in the code were retained. Domain-specific models were used, thereby gaining an additional component advantage: both the domain-specific model and the time-series data incorporated into it and encoded highly suggestively are entirely isolated from the services. This strategy improves the compartmentalisation of the platform's resources, improving security and performance.

During the later stages of CUSP platform development, the domain-specific semantic models established a connection with the domain-specific model. In addition to serving time-series data and domain-specific models, the services have become more generic. When users receive data, they must first determine how to store it in the domain-specific model before saving it. The user must review the information and domain-specific model before retrieving the data from the time series database. Although this sequence of steps observed in the user flow represented a functional approach, they were deemed to contain unnecessary overhead.

Moreover, the second distinction is that these domain-specific models are frequently protected by intellectual property. Reimplementation was required as the CUSP application moved from one domain to another. For example, this would serve as an IFC model for a water network.

Since each use case had a distinct modelling format, we had to recreate this accessing programme for each use case when testing the platform. As a result, we shifted to an open model, OWL, and, consequently, ontologies. The semantics of these components were now transferable and no longer tied to a data store or endpoint. Although a specific Application Programming Interface (API) or data format is required for a domain-specific model, users still have to query multiple data stores.

5.1.2 The balance of a mixed model approach

After careful consideration, we developed a new version of CUSP with semantics encoded in a domain-specific model. The data points were embedded in a service-specific manner and easily retrievable from the time series database despite the level of specificity of the services (implemented using a generic approach) being much lower. Due to this, we encountered significant data retrieval and preprocessing delays.

The most significant advantage was that these semantics could now be portable. At this stage, the semantic models are no longer tied to specific APIs or domain formats for a domain-specific. However, they still carried the overhead of multiple data stores).

As the CUSP platform's requirements grew in tandem with the potential real-world use case scenarios, it was divided into distinct modules, each playing the role of a microservice.

It was ultimately decided to implement a mixed-model semantic datastore to co-locate the time series data and the domain data, which includes information on diverse topics such as the buildings, the city, and the water flow path. This strategy could be confidently asserted as being feasible in light of the following factors:

- the large amounts of data that needed to be ingested by the platform
- the potentially high number of API lookups a typical user would perform during a typical usage session
- the overall smooth operation of all system components.

Although some semantics are temporarily stored on the user's local machine, they are ultimately transferred to the *Semantic Datastore*. In addition, the user cannot directly use OWL due to performance considerations.

5.1.3 Technical solutions for modelling ontologies, semantic interoperability and semi-automated model generation

The CUSP platform adopts an inclusive approach that allows all users, regardless of their technical expertise, to interact dynamically with each facet of a Digital Twin.

As explained in the previous chapters, a Digital Twin represents a virtual replica of a physical entity, system, or process, such as a building or an urban area, providing a medium for simulating, analysing, and understanding its behaviour. The platform is designed to ensure that a user, whether a city planner, building manager, or resident, can easily interact with Digital Twin, extract valuable information, and make informed decisions.

Key to this capability is the underlying semantic technology, including ontologies, that the platform employs. By modelling complex concepts and relationships within the Digital Twin, the platform provides an intuitive understanding of the virtual replica, reducing the barrier of technical jargon and enhancing the accessibility of the system to a broader range of users.

The platform uses visually annotated data structures, an approach inspired by CUSP. This allows for a user-friendly interface that presents data and relationships in an easily digestible format with various visualisation models. Consequently, a user can interact dynamically with every component of the Digital Twin, comprehending their decisions' context, interconnections, and potential impacts.

Furthermore, by fostering semantic interoperability, the platform integrates data from diverse sources, enabling seamless interaction between multiple services. This can significantly enhance the user's understanding of the system and facilitate a unified and consistent user experience.

In addition, the platform enables personalisation and adaptivity, tailoring services, and information delivery according to each user's unique needs and preferences. It supports intelligent search capabilities, providing users with more relevant and precise results, further improving the user experience.

Once the user has gained a comprehensive understanding of this complex data model, the user can draw conclusions and inferences, operations that could be accomplished using other accessible languages. As stated above, SWRL is among the primary ones. Nevertheless, as stated previously, ontologies are founded on the open-world premise. Therefore, new contributions to ontology are always possible, and machine inference must account for this.

Consequently, the user cannot draw conclusions without data. Frequently, when the user performs reasoning on ontologies, especially if unfamiliar with them, the user would receive unexpected responses. This often indicates that what the user assumed cannot be determined or that the query needs to be corrected because the user had not explicitly stated that anything would occur. Therefore, users must model all potential properties and scenarios.

A non-unique naming assumption is also present in ontologies. The user should never assume that a term in an ontology is unique or non-existent. They may be distinguishable.

In other words, according to the closed world assumption, if something is unknown to be *true*, it must be *false*. It is assumed that the world is accessible. In this situation, the name needed to be recognised as authentic. Because humans are incapable of being offensive, this close- and open-world assumption must be considered with every event of interaction with ontologies.

In particular, people trained in computer programming have the implicit assumption that the machine will assume something is *false* unless they are told otherwise. In effect, it is a multivalued logic. If users had previously created single-value logic code using an ontology they had created, they might need clarification. Moreover, if *true* and the user receives an accurate response, assume that it is *true*; otherwise, assume that it is *false*. Moreover, the unknown is lost in this manner.

Building Smart manages IFC OWL, but it is an exact copy of the IFC data model, whereas link building data ontologies are generated. There was no way to find a lighter version of IFC or a semantically lighter representation. However, the problem was within the fact that the user could easily convert from IFC to link-building data ontology, but not in the opposite direction without encountering issues. However, ontology data was often so sparse in many built-environment domains that users were forced to add their own classes and then attempt to align them with other elements as the use case evolves. The platform's underlying semantics were designed to be highly high-level, allowing users to add new semantics for novel case studies.

Building models are extensible like any other semantically represented entity in the real world. We could include zoning as part of the modelling, reasoning, or inference process, and we can define several types of zone: fire zone, walking zone, electrical zones, user zones (at the local/district level), and political zones (at a city or national level). Spatial models also incorporate the notion of systems with nodes, edges, devices, and transmission elements. The research offers an abstract version of that. Furthermore, the researcher is currently considering the actual ontologies model, which includes heat systems, electrical systems, and water networks. If the user had a different system, the semantics of that system could be easily added to the ontology by creating new subclasses of the existing classes.

However, it is extensible to model anything with a graph-like structure. An entity is the starting point here. Subclasses of entities include locations and material objects. The researcher is also making server-side operations accessible at the level of a zone, space, floor, or block.

The approach created a high-level model that merely connects the components. As new case studies were encountered, new capabilities were added incrementally. The development strategy was based on publicly accessible semantics. Modifications were required to achieve optimal performance and adaptability. Unfortunately, numerous types of ontologies adhere to the open standard for extremely abstract situations, such as sensing. Due to the highly abstract and complex nature of these scenarios, the potential for bloated data to result from their application was considered. Since bloated output would negatively impact the platform's performance, it was determined that these should ideally be reused as-is.

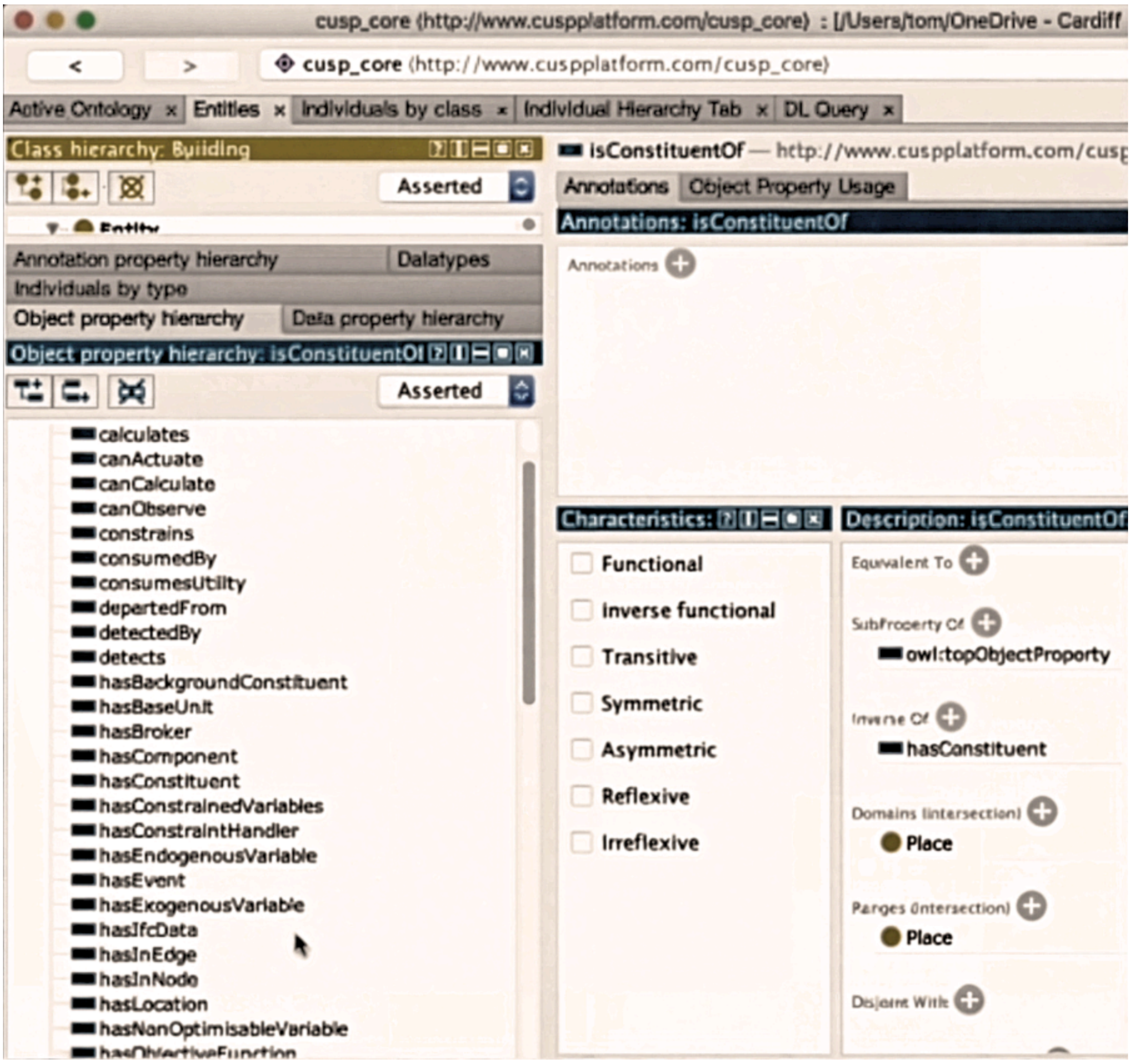


Figure 21: The CUSP ontology modelled in Protege

Fig. 21 above illustrates the main components of the CUSP ontology visualised in Protege software. The researcher has a score called components and general abstract concepts CUSP units, which enables us to model SR units and composed units (i.e., kilowatts, kilowatt-hours, metres, squared metres, etc.).

Two customer values and a CUSP time series were implemented, allowing us to represent numeric values and then arrange them in a time series. The researcher has queries that represent built-in queries.

The platform has the ability to perform custom queries, and these are represented within the ontology. The user can define custom queries in their custom ontology, which subsequently becomes available within the platform. The CUSP platform ontology also enables the sensors to model through spatial environmental modelling.

5.1.4 Stimulus components

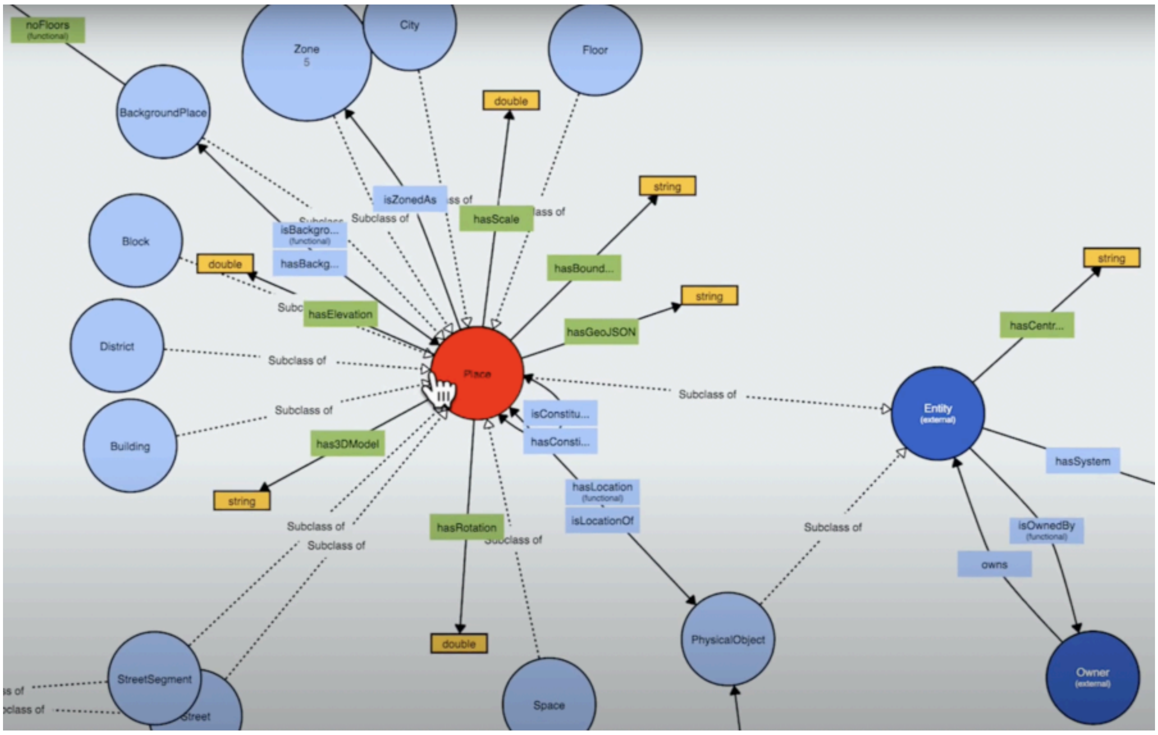


Figure 22: A capture of the WebOWL ontology visualiser

Fig. 22 an ontology visualiser called WebOWL, which features a reasonably accurate representation to the user, showing the basic structures and properties, both of which could have a unit and a value. Furthermore, webfoot is related to a semantic language-related term called a “stimulus”.

A *stimulus* is a trigger in the real world that can provide input to a sense or be affected by an actuator. An external environmental factor, such as temperature, energy usage, or maximum occupancy of people, could represent a stimulus. Some aspects could be monitored or affected in the real world. A sensor actuator, a device like a piece of hardware, detects an actuator given a stimulus as an input or a given physical location to a former property. If The user had a sensor that could monitor temperature and put that sensor in a building, The user attached it to a wall, but the user had also configured what it monitored airspace.

A sensor has a set of properties to show. It could observe one or more stimuli. A sensor is a subclass of a physical node. Furthermore, as an entity hosts it, the sensor is not showing because it only shows the sense of an aspect of the ontology in its subclasses of this entity property of types, like spaces, buildings, floors, roads, and streets. Through all these relationships, it could be noticed that a sensor is actually hosted by something tangible in the real world, observing a given stimulus featuring several properties.

The user has an actual table: *observable*, *calculated*, *predicted*, *simulated*, and *optimisable*. This approach enables the platform to differentiate between where various data points originate from a data stream. We can have quite complex structures, like a sensor that can monitor multiple parameters that could generate multiple properties. If the user had a sensor that could monitor things in multiple locations, use form properties at each location. The user might have a particular competition. It is monitored by a sensor and predicted by a AI model. The user also had two components: one optimiseable (a "toy" component) and another predicted (observable). There is also the concept of a "possible property". A "possible property" is when the semantics indicate that property should exist, but no data were present at that time. This approach enables us to deal with situations where we know that a sensor can record those data points or the researcher has information that there is a model capable of recording people producing their prediction or simulation of those data. However, the researcher has not yet encountered this situation in use cases and demonstrations.

All data points are restored in the form of event objects. These event objects give the user the ability to engage in various operations, such as calculation simulation, prediction, observation, and actuation. Moreover, events are inherently linked to properties, whereas properties are reciprocally associated with events. Consequently, the user has the ability to discern the specific property that originated from a given event.

Additionally, users have the ability to determine the correlation between events and specific properties. Additionally, the user is empowered to retrieve all events associated with a particular property, such as all events pertaining to a specific park within a defined time frame. Alternatively, the user can employ a reverse approach by searching for all properties linked to a specific type of event. The direction of the query depends on the ontology's efficiency and response time, which are determined through benchmark evaluations.

The user's trim-down process is tailored according to their individual preferences. Even with the inclusion of indexes, this operation is executed with notable efficiency. Therefore, the decision to execute the query in reverse is motivated by the speed that it offers.

A significant number of platform features are designed and implemented with a focus on speed optimisation. As mentioned previously, one rationale for adopting a layered approach is to compel users to efficiently conduct queries, thereby validating them through the web user interface rather than permitting inefficient manual queries directly on the back-end database.

5.1.5 Tree of classes and sub-classes

Within the tree-like ontological representation of a web application, the concept hierarchy is structured through the presence of *classes* and *subclasses*, forming a layered taxonomy. The hierarchy within the ontological representation enables the systematic organisation and categorisation of entities based on their inherent characteristics and relationships. The researcher observes that at the highest level of this hierarchy, the broad category of "buildings" serves as the parent node, encompassing various structures within the built environment and providing a comprehensive context for more specific types of building.

As the researcher moves down the hierarchy, subclasses emerge that represent different types of building, including hospitals, airports, and schools. Each subclass functions as a child node, inheriting the attributes and characteristics of the parent class, while introducing unique attributes specific to its respective building type. For example, the hospital subclass may include attributes related to the availability of medical facilities and specialised departments.

Further down the hierarchy, the researcher encounters the "street" class, which embodies a broader concept associated with transportation and urban infrastructure. Serving as a parent node, this class offers a higher level of abstraction within the ontology and is subsequently followed by subclasses representing specific segments of the streets.

These "street segment" subclasses inherit the attributes and properties of the "street" class while introducing segment-specific characteristics, such as landmarks or specific traffic regulations.

By adopting such a hierarchical structure within the ontological representation of the web application, the researcher acknowledges that the entities are logically organised, facilitating efficient data management, retrieval, and analysis. This structure establishes a clear taxonomy that aids in navigating and comprehending the relationships between different entities within the system, enabling more effective decision-making and facilitating application development.

The streets are generally divided into street segments, as CUSP platform users, as some streets can be surprisingly long. From a technical standpoint, users have the ability to model streets in unconventional ways (e.g., creating a single long road spanning from Cardiff to a point west of Swansea). However, typically, users are not motivated to reason at such granular levels, but rather to divide lengthy streets into smaller segments.

During demonstrations and exploration of use cases, the CUSP platform incorporated concepts of district and city. It is important to note that the researcher has not yet progressed with a large-scale model at the national level.

Users have the ability to effortlessly introduce new subclasses and establish new relationships, which are essential inputs for CUSP algorithms to generate larger entities. However, this process may require some laborious effort on the part of the user, including explicitly specifying relationships derived from data-driven modelling. These relationships predict values based on other values within the system, thus training the underlying middleware to determine when to execute these prediction models.

In addition, the platform allows users to model the performance of various components. Users may possess multiple data-driven models for a given property, and there may be no predefined rules for specific time intervals. However, new rules can easily be created during the use case development lifecycle based on periodically derived data from data-driven models.

When new data are obtained, users, assisted by the middleware, have the option to selectively run the model with the highest accuracy. Alternatively, users can choose to run all models simultaneously and filter the results. These capabilities are facilitated by the ontology, which models the information of all these software components. Furthermore, users can conduct custom queries more effectively than the ontology itself and display the query results on the web interface via a REST (Representational State Transfer) API.

In the ontological representation of the researcher, there are building blocks, city or district floor spaces, street segments, and individual entities. Note that all the available information is retrieved from the super-class, the parent class, and the explicitly modelled entities. These entities explicitly define the properties of real-world assets, for example:

- each building has a name which is a string
- each model has some centre coordinates
- each model might have a Message Broker or some boundary coordinates associated with it
- each building might have constituent property.

The user occasionally has visibility into a subset of properties that can act upon it in this context, hence the utilisation of web-based visualisation. Additionally, the user has access to all the properties present in the ontology, including hierarchical relationships such as:

- a place linking to another place
- a space is on a floor
- a floor is in a building
- a building is on the street
- the street is in a district
- the district is in the city

The user can define custom hierarchies to establish connections between objects. The only information that can be provided on this particular property is its domain during rainy conditions.

Moreover, there exists an inverse relationship that is not a constituent of this property, similar to the primary relationship, but the user has obtained the "house" constituent. This inverse relationship represents a distinct counterpart. These properties are employed in various sections of the ontology.

To enhance usability, the web-based viewer offers a user-friendly alternative to the ontology tool, requiring less technical expertise for model editing.

The Explorer tool displays all properties that can be narrowed down on the basis of the active ontology or filter entities related explicitly to spatial aspects. However, it should be noted that this approach is not flawless, as it still includes some aspects such as the ”broker” from the algorithm ontology and the IFC route from IFC OWL. Therefore, it does not completely restrict the inclusion of these elements.

5.2 The CUSP platform underlying Technical Infrastructure

This section details the various APIs and how their underlying services and User Interface components interact with the bespoke middleware components.

Each component of the CUSP platform achieves a specific goal through the continuous performance of a specific associated task. In several instances, the platform encounters additional computational overhead. Data stored in the platform and trigger data science services are required, which are part of the conceptual architecture of the platform. The various CUSP services interact with the User Interface through the actual middleware components. A service achieves a specific computational task. The various UI components are an interface to the backend through the APIs to access the data stored in the platform and subsequently trigger the data science services as required.

Data science services were developed by other members of the research group. The researcher was only responsible with development of the user interface.

The CUSP middleware embeds several components that form a coherent set of scientific and

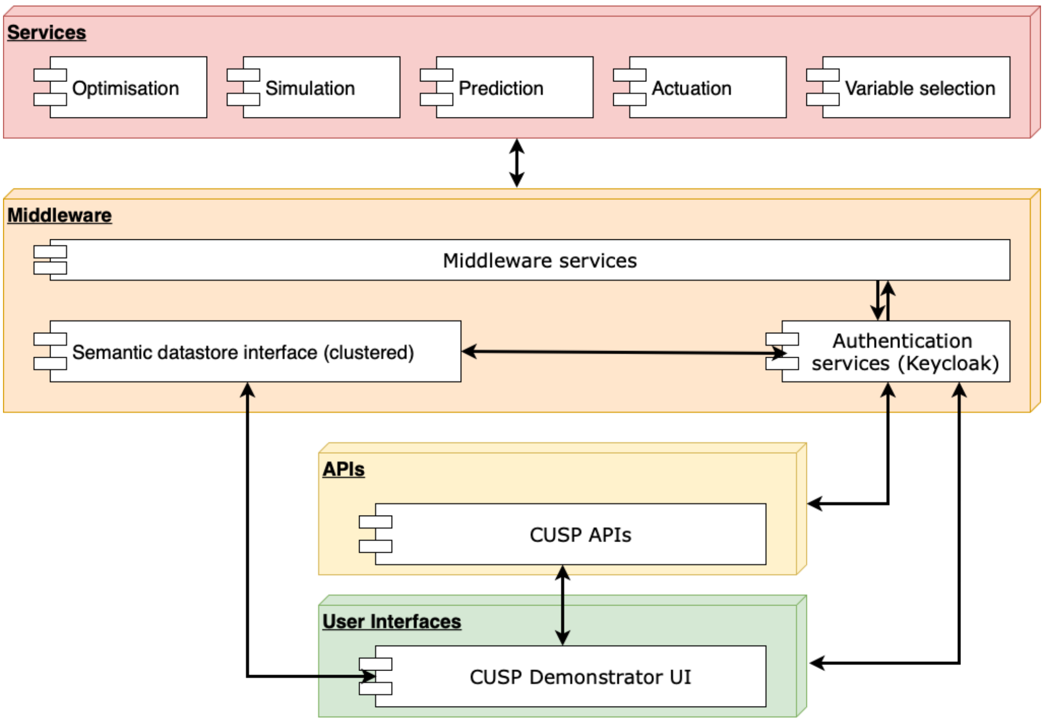


Figure 23: A Digital Twin ontology modelling approach

analytical tools, such as *linked data*, *simulation*, *optimisation*, and *prediction* components.

The CUSP modules/agents are "programs that operate at a semantic level high enough to form new connections with other programs to perform a task". Functional agents are powered by decision networks that perform one or more of the following functions: making binary decisions (when to give YES / NO permission to execute a particular function or operation), when to trigger a specific action / appropriate response, and deciding on the most appropriate model to run, based on *semantic similarity*.

This CUSP middleware is accessible through dedicated APIs. A tailored version of the middleware contains:

1. a Semantic Data-Store

2. a security component (authentication and security services)

3. a semantic middleware

4. data science services (including models for *optimisation*, *prediction*, *simulation*, and *actuation*).

5.2.1 Authentication / Security services

The researcher had no intention of reinventing well-established design patterns when it comes to security, as there are readily available scalable open source solutions for authorisation and authentication services. To address this, the researcher uses Keycloak, an open source security and authentication library that allows seamless integration with third-party REST APIs, such as Office365 or custom-developed APIs. In addition, the researcher has developed an Integrator component to facilitate the distribution of security-related messages across the application, ensuring effective communication with authentication services and other data components.

5.2.2 Database Back-End

CUSP uses a graph-based database backend to store data (the semantic datastore). This internal component is not accessed directly by any services exposed to the user. This architectural decision has several arguments in favour of enhancing the user experience.

5.2.3 Data Abstraction

CUSP promotes data abstraction by shielding the user from direct access to the semantic datastore. This abstraction layer allows the user to interact with the system through more user-friendly and intuitive interfaces. Instead of dealing with the complexities and intricacies of the database structure, the user can focus on performing their tasks and achieving their goals efficiently.

5.2.4 Simplified User Interaction

Direct access to the back-end database could introduce technical complexities and challenges for users who may need to gain in-depth knowledge of database management systems. By abstracting the database through intermediary services, CUSP simplifies user interaction, reducing the learning curve, and allowing users to perform their tasks more effortlessly.

5.2.5 Enhanced Security

The user experience encompasses not only usability but also security. By isolating the semantic datastore from direct user access, CUSP improves the overall security of the system. This approach ensures that sensitive data and underlying database structures are protected from potential malicious activities or user-intentional alterations, providing a controlled and secure environment for data storage and management.

5.2.6 Performance Optimisation

By decoupling user-facing services from direct access to the semantic datastore, CUSP can implement performance optimisations tailored to specific user interactions. The system can leverage caching mechanisms, query optimisation, and other techniques to improve response times and provide a seamless user experience. This enables users to retrieve and manipulate data efficiently, contributing to overall user satisfaction.

5.2.7 Scalability and Flexibility

Separation between user-facing services and the semantic datastore enables CUSP to scale and evolve its back-end infrastructure without impacting the user experience. As the system grows and handles larger data volumes or introduces new features, the underlying database infrastructure can be modified or upgraded, ensuring minimal disruption to the user experience. This flexibility allows CUSP to adapt to changing requirements and technological advancements while maintaining a consistent and reliable user experience.

5.2.8 Data Science Services

The components of AI (data science services) of the CUSP platform include the following:

1. an *objective function*
2. a *feature selection service*
3. a *training service*
4. a *prediction service*

5. an *optimisation service*

The primary data science services (Optimisation, Prediction, Simulation) offer the user the ability to construct bespoke AI models. A set of Model Construction services for these three scopes is deployed and triggered by the User Interface or the Middleware. These data science services retrieve events from the semantic data store as and when required. Depending on the complexity of the Digital Twin scenario and the "problem site", many optimisation services can be deployed for different problems, and the AI models are automatically built. The user can duplicate them and run them in parallel for different problems. A brief overview of the CUSP The data science workflow of interaction between a user and the CUSP platform is shown below.

1. the *objective function* receives, from a user, an optimisation objective associated with a built environment
2. the *feature selection service* performs semantic problem interpretation based on the optimisation objective
3. the *optimisation service* semantic optimisation for the optimisation objective
4. through the Graphical User Interface (GUI), CUSP provides an optimisation output to the user
5. the *training service* receives a semantic domain model of a built environment
6. through the GUI, based on the semantic domain model of the built environment, CUSP renders a three-dimensional model of the built environment
7. through the GUI, CUSP renders a consumption graph comprising consumption data related to the selected element of the three-dimensional model
8. through the GUI, CUSP receives through the *objective function* an optimisation scenario comprising one or more optimisation features in a natural language format
9. the *training service*, based on the optimisation scenario, one or more machine learning models
10. the *feature selection service* is an optimisation algorithm based on one or more machine learning models
11. through the Graphical User Interface (GUI), an optimisation output is provided

This functionality brings several benefits to the user experience.

5.2.9 Customisable AI Models

By allowing users to construct their own AI models, CUSP empowers them to tailor the models to their specific needs and problem domains. This level of customisation enhances the user experience by enabling users to address their unique challenges and obtain more accurate and relevant insights.

5.2.10 Model Construction Services

CUSP deploys a set of Model Construction services that facilitate the creation of AI models for Optimisation, Prediction, and Simulation. These services are triggered by the User Interface or the Middleware, making it convenient for users to initiate the model construction process. This streamlined approach simplifies the user experience and reduces the barrier to entry to use AI capabilities.

5.2.11 Seamless Data Retrieval

The data science services retrieve events from the semantic data store as and when required. This seamless data retrieval ensures that AI models are always up-to-date with the latest information, improving the accuracy and relevance of predictions and optimisations. As a result, users benefit from a more reliable and trustworthy user experience.

5.2.12 Scalability and Parallel Execution

Depending on the complexity of the Digital Twin scenario and the specific problem site, CUSP allows the deployment of multiple optimisation services for different problems. Users can duplicate and run AI models in parallel, addressing various challenges simultaneously. This scalability and parallel execution capability enhance productivity and efficiency, enabling users to achieve faster results and explore multiple scenarios, ultimately improving the user experience.

5.2.13 Automation of AI Model Building

CUSP automates the process of building AI models, reducing the manual effort required from users. With the ability to automatically build AI models based on the complexity of the Digital Twin scenario, users can focus more on analysing results and making informed decisions rather than spending time on manual model construction. This automation simplifies the user experience and accelerates the adoption of AI technology.

5.2.14 Connectors

To connect the CUSP platform to external data input sources, a connector was implemented, which links the interfaces to the underlying native APIs and the data source they consume. Therefore, the connectors also inherited the role of producing CUSP events as data points are received.

5.2.15 The Semantic Data-store

The Semantic Data-store is the REST web service interface to the data stored within the platform. It can be accessed via a REST web service interface to the data stored by the platform, implemented in the Java programming language, on top of the Tomcat web server. APIs can be used to inject or retrieve events from the Message Bus.

The *OrientDB database* is the revised back-end database for the current use case, and we can hold multiple use cases per building being instant. In that sense, DB has multiple databases or tents, so they are entirely separable. The database can also be distributed and split. However, the researcher has not yet explored this because the researcher has not yet shown a use case beyond what the existing DB used in conjunction with one High-Performance Computing (HPC) node. A more efficient way to scale it is to separate each instance onto a single node and only when one instance is getting so big that the user needs more storage.

There are unresolved controversies about sharding between the industry and the academic community. However, peer-reviewed articles and benchmarking performance data show encouraging results achieved as a result of implementing this technique. Therefore, it was decided that a graph-based database was used due to performance considerations. Generally, all the storage instances that are based on ontologies natively tend to exhibit relatively low performance levels compared to their counterparts implementing sharding. For example, we mentioned earlier the need to query all sensor data centre events on a given building.

If we implement this novel database, we must search across all event repositories and filter by building a graph database. Even in the given scenario, indexing on a relational database enables us first to select the building and then efficiently locate only the associated events.

Examining the performance literature of graph databases as opposed to relational databases reveals that this second approach is significantly more efficient. Typically, a relational database outperforms a graph database if the user desires to retrieve a large proportion of the graph.

However, in the context of a built environment use case, this is rarely desired. A graph database would outperform a relational database by a wide margin even in the most marginal use cases.

The performance of different flavours of graph databases remains a challenging question. OrientDB appears to be highly efficient compared to other database management systems and query languages, exhibiting at least "on par" with some of its counterparts.

Additionally, there is nothing to suggest that the platform could not be modified to support an alternative RDBMS (Relational Database Management System), which is part of the standard OrientDB suite (which CUSP is currently using). Theoretically, these transitions would not require any code or configuration parameter modification. However, without a clear and robust justification for migration, this project will be put on hold until further notice.

The OrientDB back-end enables users to perform manual queries on the data. From this, it did a very similar job to a typical SQL database, like in the "LIKE" query (looking for values similar to the defined query arguments).

The semantic ontology can expose data for some properties, such as predicted calculated properties, which are properties calculated for a prediction operation. We can also inspect the link to the algorithm that calculated them, highlighting the different data set parameters used to calculate the prediction results. We can also inspect the prediction timestamps and the latest values.

5.2.16 The REST APIs

Platform APIs send and receive events from the ***Message Bus*** and access data from services using ***the Orchestrator***. Two APIs (Java and Javascript) have been developed to provide an interface at a higher level of abstraction than the REST / MQTT interface, namely:

- Data service APIs are responsible for the reciprocal communication bus between Data Science Services and the middleware.
- Consumer APIs, featuring the ability to ingest large data models (since the platform can consume various types of data: GIS data, structure data about a problem site into the platform).

Through its REST API, the CUSP architecture enables a loose coupling between the AI / Data Science Services and the simulation services. Therefore, the bare minimum essential task that a newly created AI / Data Science or simulation service needs is to respond to calls through the simulation or AI, depending on whichever route the service is linked with through the API.

Additionally, a newly created service does not get throttled down by querying the Semantic Datastore. The user could choose to retrieve manually should they wish to allocate resources to building new modules within the CUSP platform, which eventually feature more complex functionality. The base data set for each service is always delivered by the Orchestrator, rendering them to load in developing new AI models utilising the Data Science Services, such as Prediction or Optimisation. The Services module is lighter in terms of computational requirements than the Orchestrator.

The Sandcastle JavaScript APIs can send methods and messages on the event bus, which constitutively wraps a standard JavaScript MQTT API. We also have the custom main API that had been implemented in Java, with the dedicated functionality of wrapping a Java MQTT library. The complete functionality of the REST web services is to retrieve and process data in a production Smart City environment.

The figure above illustrates the root of the REST Web services API, of which the output is entirely visualisable in the JSON format, similar to a configuration set. REST API provides the user with valuable and comprehensive information on all visible entities in the User Interface. Suppose that we navigate to the root of the ontology representation (the head node representing the primary parent "City of Cardiff" node). In that case, the REST API ingests a significant quantity of data for the City of Cardiff, including, but not limited to, building properties and their names, their hosts (infrastructure and street segments) and the rest of the constituent elements of the city.

Every building on a map generated by the CUSP platform is represented as a unique URI (Uniform Resource Identifier) to promote accountability in designing and interacting with the ontology and data integrity. In the figure above, the system communicates that this view is missing in navigation through the spatial structure.

5.2.17 Historical Data and Time-Series

The CUSP platform allows queries based on historical data. If the user zooms in on this bit here, we will send this query as a query parameter comprised of:

- a start absolute variable value and an end absolute value represented by timestamps
- the user can also perform a relative query, querying for operations started/completed at set time intervals (i.e., a simulation that started ten days ago and was completed three days ago).

The query only returns a value and a timestamp, which APIs or other software tools would then parse to generate graphs or perform further analysis. Users have the ability to perform start, relative, and absolute queries. Using a relative approach, the user can say one week, two weeks, or one year. The user can limit the number of results that are retained.

By decoding and adding aggregators, the user can perform tasks such as counting the number of results, average, sum, last Max, and Min-first.

The server is taking the weekly average over this data set to speed up the data retrieval process and efficiently cache the responses of frequently executed queries. Therefore, at every query, the server will return a weekly average instead of the raw data - this is a custom component developed in Java and distributed as a Docker container.

In the background, data are transmitted to the platform. The message picked up first is distributed to the other components of the Orchestrator. Arrived here, the Orchestrator decides on the appropriate course of action and either passes it further, stores it in the historical dataset, or merely disposes of it. The Synchroniser will also pick it up and update the database. When retrieving the data, the API puzzles together all the data fetched from the back-end, compiling it for the user in a readable format.

5.3 CUSP User Interface Contexts and Components

The construction of the User Interface (UI) and the grouping of components into specific contexts within the CUSP software is a testament to thoughtful design and careful consideration of user needs and usability principles. This process involves understanding the tasks that users need to perform, the information they need to access, and the actions they need to take. The UI components are then designed and grouped into contexts to support these tasks, provide the necessary information, and facilitate the required actions.

5.3.1 The Authentication Context

The Authentication Context is designed to provide a secure and straightforward way for users to log into the system. The authentication screen is a standalone UI component that serves a single purpose: user authentication, focusing on a single task within a context simplifies the user experience and reducing cognitive load.

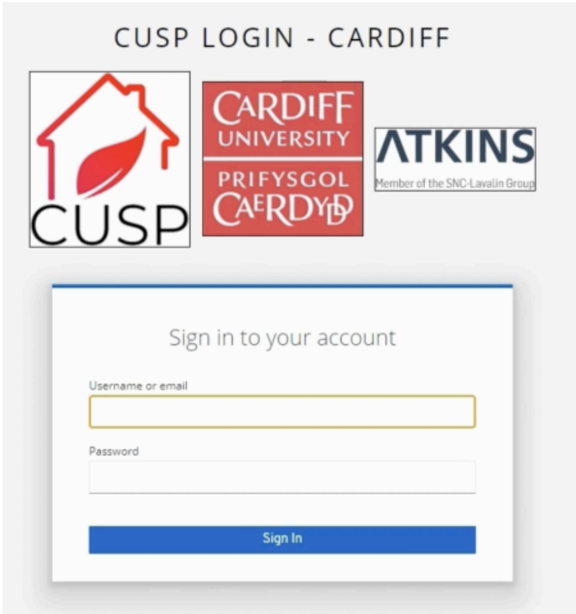


Figure 24: The CUSP authentication context

5.3.2 The Building and World Visualisation Context

The Building Visualisation Context is designed to support viewing 3D models of buildings. The UI components in this context, including the building loading progress indicator and the viewer surface, are specifically designed to facilitate this task. The viewer surface supports panning, zooming, and orbiting, providing users with a range of tools to interact with the 3D models in a way that enables the visualisation of a model from all possible angles.

The World Visualisation Context is designed to support the task of visualising the location of a model on a global scale. The World Visualiser is a UI component that pinpoints the location of a model, providing users with a broader context for their data.

5.3.3 The Task Selection Context

The Task Selection Context is designed to facilitate the selection of different functionalities within the software. The main taskbar is an UI component that provides a consistent and easily accessible way for users to switch between different tasks such as *Trends*, *Events*, *Predictions*, *Actuations*, *Optimisations*, and *Composition*. This consistent access to critical functionalities across different contexts enhances usability by reducing the number of steps required to perform tasks.

Figure 25: The CUSP world visualisation context

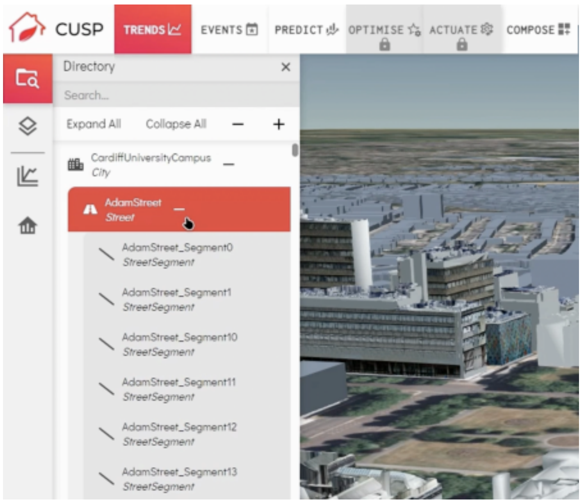


Figure 26: The CUSP entity selection context

5.3.4 The Entity Selection Context

The **Entity Selection Context** within the CUSP software is designed to facilitate the task of selecting entities that are integral components of the semantic ontology. This context is crucial, as it allows users to interact with the system's underlying data structure and choose specific entities for further analysis or manipulation.

In this context, the primary User Interface (UI) component is the Directory in the left taskbar. The Directory presents a hierarchical view of the entities based on their relative position within the semantic ontology. This hierarchical structure provides users with a clear understanding of the relationships between different entities, aiding in navigation and selection.

The Directory is designed with several features to enhance usability. For example, if an entity has at least one child node, a "plus" icon appears next to its name. Clicking on this icon expands the entity to reveal its child nodes, allowing users to navigate deeper into the hierarchy. Conversely, a "minus" icon appears next to an expanded entity, which users can click to collapse the entity and hide its child nodes. These interactive features make the Directory dynamic and easy to navigate, enhancing the user experience.

Another feature of the Directory is the navigation buttons located at the top of the window. These buttons help users navigate through the Directory, providing additional navigation options. This feature improves usability by offering multiple navigation methods, catering to different user preferences and work styles. The Directory also includes plus and minus buttons on the right side of the component, which allow users to adjust the viewing scale according to their preference. This feature gives users control over the level of detail they want to view, enhancing usability by catering to individual user needs and preferences.

In addition to the Directory, the Entity Selection Context includes a feature allowing users to select entities by clicking on a 3D model within the viewer's surface. This feature provides a more intuitive and interactive method for selecting entities, particularly for users who prefer a visual approach to data interaction.

5.3.5 The Layer Manipulation Context

The Layer Manipulation Context within the CUSP software is designed to facilitate the task of controlling the display of low- and high-resolution models. This context is crucial as it allows users to manage the level of detail they view within the 3D building models, providing a balance between comprehensive detail and overall perspective.

The primary User Interface (UI) component in this context is the Layers tab. The Layers tab is a dedicated UI component that allows users to toggle between the display of low-resolution and high-resolution models within the viewer. This feature gives users control over the level of detail they want to view, enhancing usability by catering to individual user needs and preferences.

The high-resolution models are broken down into multiple levels: buildings, blocks, floors, zones, and spaces. This hierarchical structure provides users with a clear understanding of the relationships

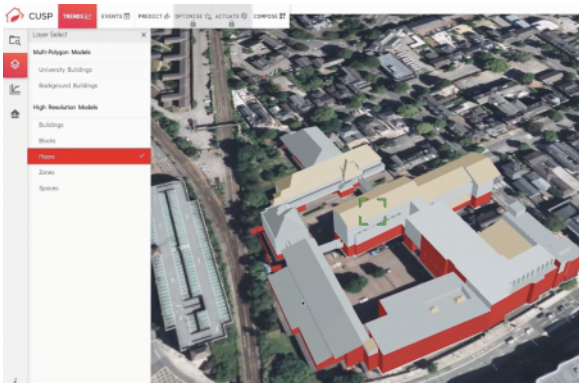


Figure 27: The CUSP layer manipulation context

between different parts of the building. It allows them to navigate through the building structure in a detailed and systematic way.

However, only one of these levels can be visible at one time. This design decision prevents the viewer from becoming cluttered with too much information, which could overwhelm the user and hinder their ability to interpret the data. The software improves clarity and reduces cognitive load by allowing users to focus on one level at a time.

The Layers tab also provides a consistent and easily accessible way for users to switch between different levels of detail. This consistency improves usability by reducing the number of steps required to perform this task and making the process more intuitive.

In summary, the Layer Manipulation Context is designed with various UI components and features to control the display of low- and high-resolution models. By providing a clear and intuitive Layers tab, a hierarchical structure of high-resolution models, and a consistent method for switching between different levels of detail, this context enhances usability and provides a positive user experience.

5.3.6 The Information Viewing Context

The Information Viewing Context within the CUSP software is designed to provide users with additional information about the selected building. This context is crucial, as it allows users to understand the entities with which they interact, enhancing their ability to make informed decisions and perform detailed analyses.

The primary component of the user interface (UI) in this context is the information button. When clicked, this button triggers the display of additional information about the selected building. The information provided includes details about the building, such as its coordinates, the data types available within the ontology, and associated assets, such as IFC files and Energy Plus models. This comprehensive information enhances the user's understanding of the building and provides valuable context for their analyses.

The Information button is designed to be easily accessible and intuitive to use. Its placement in

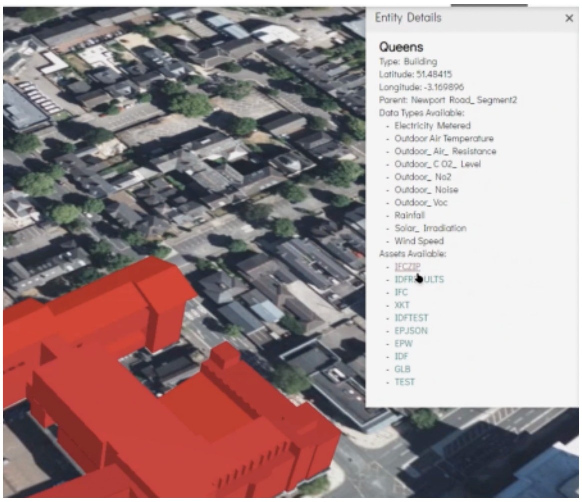


Figure 28: The CUSP information viewing context

the upper right corner of the interface ensures that it is consistently available to users, regardless of their other tasks. This consistent access to the Information button enhances usability by reducing the number of steps required to view additional information about the building.

The display of information triggered by the Information button is designed to be straightforward to understand. Information is presented in a structured format, with different types of information grouped for the sake of clarity. This structured presentation of information reduces cognitive load and makes it easier for users to find the specific information they need.

In addition to providing information about the building, the Information Viewing Context also facilitates downloading attached files to the selected entity. This functionality improves usability by giving users easy access to essential building assets.

5.3.7 The Graph Generation Context

The Graph Generation Context within the CUSP software is designed to facilitate generating and interacting with graphs. This context is crucial, as it allows users to visually visualise the data in a meaningful way, enabling them to gain insight and make informed decisions.

The primary User Interface (UI) component in this context is the graphs component. This component allows users to generate a graph since an entity is selected. The graph provides a visual representation of the data associated with the selected entity, making it easier for users to understand and interpret it.

The Graphs component is designed to be interactive, allowing users to manipulate the graph in various ways. For example, users can maximise the graph by clicking the button in the upper right-hand corner. This feature allows users to focus on the graph and view it in more detail, improving usability by catering to individual user needs and preferences.

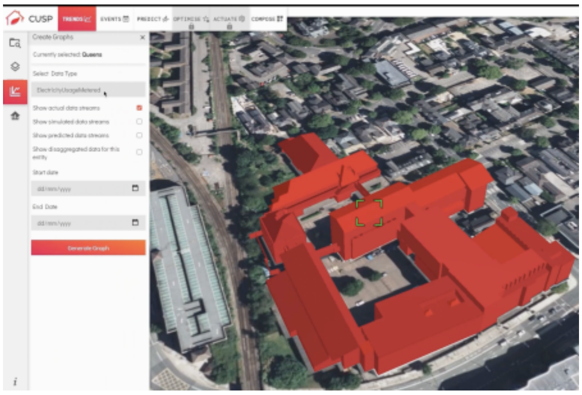


Figure 29: The Graph generation context

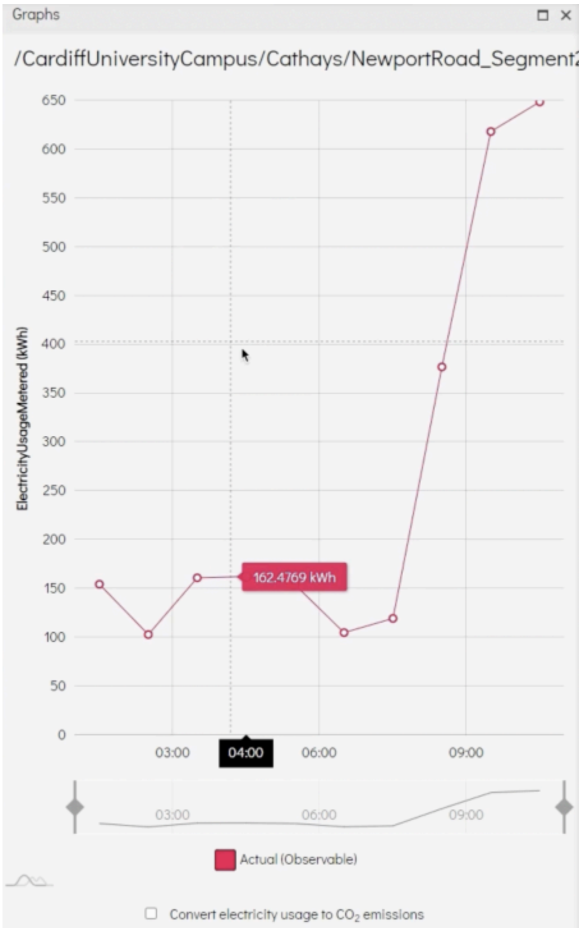


Figure 30: The Graphs Component

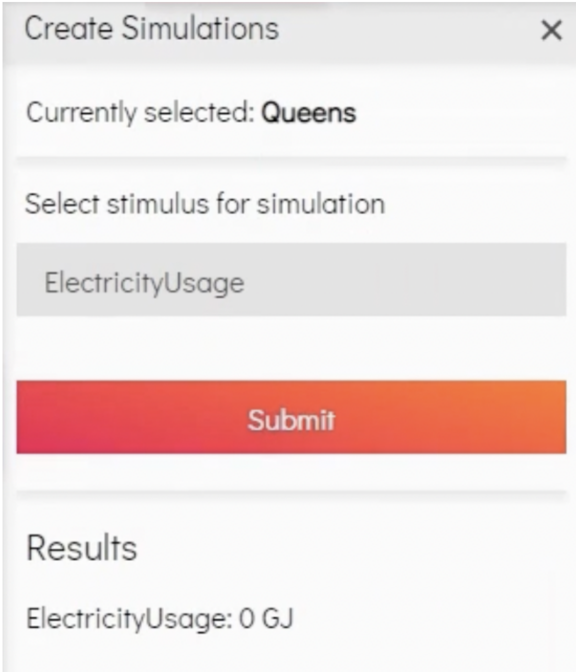


Figure 31: The CUSP simulation creation context

The Graphs component also includes various controls for interacting with the graph. These controls include the ability to zoom in on specific areas of the graph, adjust the focus on the timeline, and reset the view of the graph. These interactive features make the Graphs component dynamic and engaging, improving the user experience.

In addition to these controls, the Graphs component includes a legend that allows users to show or hide data lines. This feature gives users control over the information they want to view, improving usability by catering to individual user needs and preferences.

Another feature of the Graphs component is the data stream highlight. If the user hovers over a point on the data stream, the value at the corresponding timestamp on the x-axis will be highlighted. This feature provides immediate feedback to the user, improving usability by making the graph more interactive and informative.

The Graph Generation Context also includes a disaggregated data graph that enables users to inspect individual child entities' data contributions. This feature provides a more detailed view of the data, allowing users to gain deeper insight into the contributions of individual entities.

5.3.8 The Simulation Creation Context

The Simulation Creation Context is designed to facilitate the running of an Energy Plus simulation. The UI components in this context are specifically designed to support this task, providing users with the tools they need to create and run simulations.

5.3.9 The IFC File Import and Asset Uploader Contexts

The IFC File Import Context within the CUSP software is designed to facilitate the import of an Industry Foundation Classes (IFC) file into the semantic ontology. This context is crucial as it allows users to integrate external building data into the system, enhancing the software's utility and versatility.

In this context, the primary User Interface (UI) component is the IFC uploader form. This form guides users through the import of an IFC file, providing a clear and structured method to perform this task. The form prompts users to customise values such as building name, parent name, latitude, longitude, rotation angle, and file selection. These prompts ensure that users provide all the necessary information for the import process, reducing the likelihood of errors and enhancing usability.

The *IFC uploader* form is designed to be intuitive and easy to use. The form fields are clearly labelled, and the layout follows a logical order that aligns with the user's workflow. This intuitive design reduces the cognitive load and makes the import process more efficient.

The *IFC uploader* form also includes a feature that allows users to select the IFC file from their local system. This feature provides a familiar and straightforward method for file selection, improving usability by catering to user expectations and work habits.

Once the user has selected the file and filled in the necessary information, they can start the import process by clicking the upload button. This action triggers the file upload process, providing immediate feedback to the user and keeping them informed about the status of their request.

The *Asset Uploader Context* is designed to facilitate the uploading of an asset to an entity. The *Asset Uploader* component and the *Asset Uploader* form are UI components designed to guide users through this process. These components provide a clear and consistent method for uploading files, reducing the likelihood of errors, and improving usability.

5.3.10 The Prediction Scenario Context

The *Prediction Scenario Context* within the CUSP software is designed to facilitate the task of generating a prediction scenario. This context is crucial because it allows users to simulate different scenarios and understand their potential impacts, enabling them to make informed decisions.

The primary User Interface (UI) component in this context is the feature that allows users to generate a prediction scenario. Once an available scenario is selected, the scenario settings are displayed, providing users with a clear understanding of the current configuration of the prediction scenario. This visibility of scenario settings improves usability by informing users about the status of the system and helping them understand the potential impact of their decisions.

The *Prediction Scenario Context* also includes a feature that displays the net difference resulting from the prediction scenario. This feature, known as the cost-benefit ratio of the scenario, provides

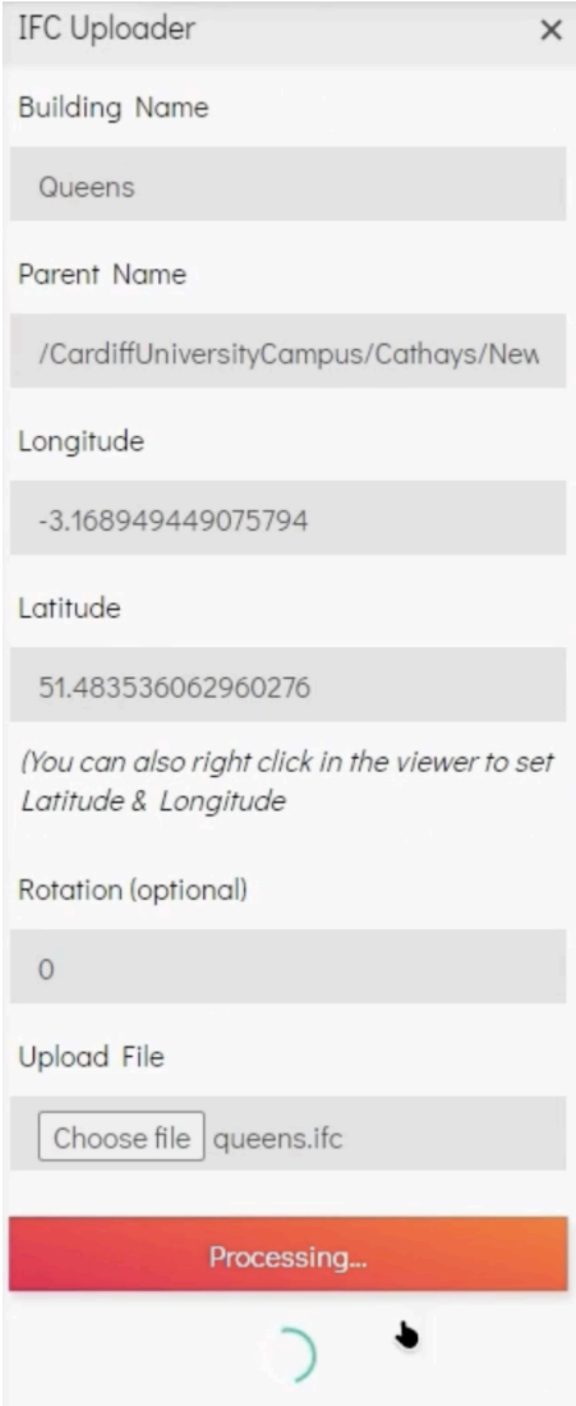


Figure 32: The CUSP IFC File Upload and Asset Uploader Contexts

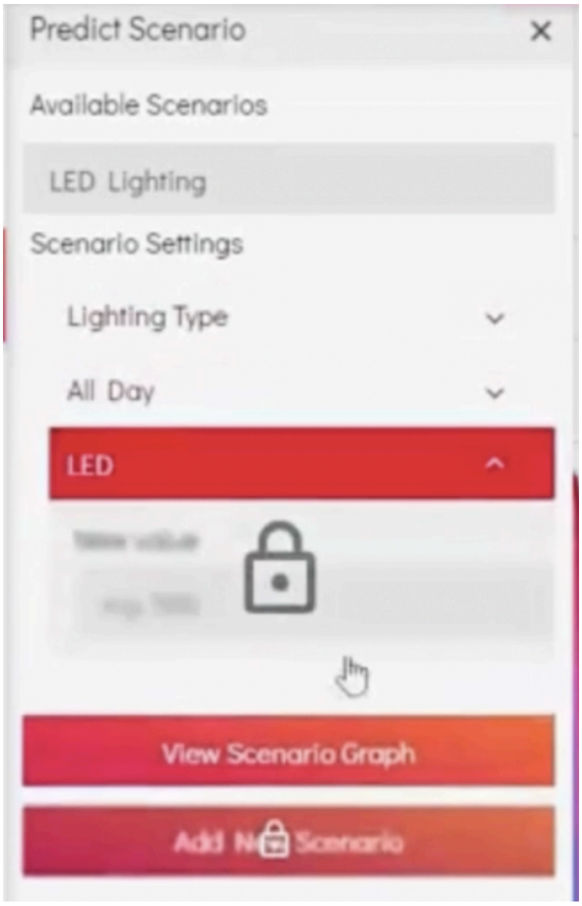


Figure 33: The CUSP Prediction Scenario context

users with a clear understanding of the potential benefits and drawbacks of the scenario, helping in decision making.

Another feature of this context is the lock pads. These are used to hide software features that are still in development. By hiding these features, the software prevents users from accessing incomplete or potentially unstable features, enhancing usability by reducing the likelihood of errors and confusion.

The *Prediction Scenario Context* also includes a *carbon emissions* checkbox. This feature allows users to visualise the corresponding carbon emissions for a scenario. By providing this information, the software enables users to consider the environmental impact of their decisions, adding a layer of complexity to the decision-making process.

The design of the User Interface (UI) components and their grouping into contexts in the CUSP software is a clear example of the thoughtful application of User Experience (UX) principles and sound usability principles. This design approach aligns with the principles enunciated by Hassenzahl, Mahlke, and Karapanos, as well as those by Norman, which emphasise the importance of creating a user-centred design that is intuitive, efficient, and enjoyable.

5.4 Adopted User-Centred Design Principles

This section re-evaluates the principles of User-Centric Design as delineated by Norman, in conjunction with the principles of effective usability propounded by Hassenzahl, Mahlke, and Karapanos. This exploration facilitates a comprehensive rationalisation for the decisions made regarding User Interface Contexts and their corresponding components, connecting them to the theoretical frameworks in the Literature Review section.

The first principle of user-centred design is understanding and addressing the needs and abilities of the user. CUSP software does this by providing a clear and intuitive interface. For example, the Authentication Context is designed to be straightforward to use, reducing the cognitive load on the user. The Building Visualisation Context provides a 3D viewer surface that supports panning, zooming, and orbiting, allowing users to interact with the buildings naturally and intuitively.

User-Centred Design (UCD) is a philosophy that places the user at the centre of the design process. It involves understanding the user's needs, abilities, and behaviour and using this understanding to guide the design of the interface. The CUSP software embodies this philosophy in designing the User Interface (UI) components and their grouping into contexts.

5.4.1 Understanding the User

The first step in UCD is understanding the user. In the case of CUSP, users are likely to be individuals or teams involved in energy management and carbon reduction. These users need to interact with complex data and perform various tasks such as viewing building models, selecting entities, generating graphs, running simulations, importing files, predicting scenarios, and uploading assets.

5.4.2 Addressing User Needs and Abilities

The UI components of CUSP are designed to address these user needs. For example, the authentication context provides a straightforward and secure way for users to log into the system. The Building Visualisation Context provides a 3D viewer surface that supports panning, zooming, and orbiting, allowing users to interact with the buildings in a way that aligns with their spatial understanding and navigation abilities.

The **task selection context** is another example of user-centred design. The main taskbar at the top of the screen allows users to easily switch between different tasks such as *Trends*, *Events*, *Prediction*, *Actuation*, *Optimisation*, and *Composition*. This design reduces the number of steps required to switch tasks, increasing efficiency and reducing cognitive load.

The **Entity Selection Context** is designed to help users select entities based on their relative position within the semantic ontology. The Directory in the left taskbar and the entity selection feature (triggered by clicking on a 3D model or entity name in the *Directory*) make it easy for users to find and select entities, reducing the time and effort required for this task.

The **Layer Manipulation Context** allows users to toggle between displaying low- and high-resolution models in the Layers tab. This feature gives users control over the level of detail they want to view, catering to their individual preferences and task requirements.

The **Information Viewing Context** provides additional information about the selected building, including coordinates, data types available within the ontology, and associated assets such as IFC files and Energy Plus models. This feature addresses the user's need for detailed information about the buildings they are managing.

Therefore, the design of the User Interface components and their grouping into contexts in the CUSP software is a clear embodiment of user-centred design (UCD). The design is tailored to the needs and abilities of the users, making the software efficient, effective, and enjoyable to use.

5.4.3 Efficiency and Effectiveness

Norman's principles of usability emphasise efficiency and effectiveness. The Task Selection Context in the CUSP software is an excellent example. The main taskbar allows users to easily switch between tasks such as *Trends*, *Events*, *Prediction*, *Actuation*, *Optimisation*, and *Composition*. This design reduces the number of steps required to switch tasks, increasing efficiency. The Entity Selection Context also follows this principle by allowing users to select entities from a directory, reducing the time and effort required to find and select an entity.

As Don Norman emphasises, efficiency and effectiveness are fundamental principles of good usability. An efficient interface allows users to perform their tasks quickly, while an effective interface ensures that tasks are completed accurately and completely. The design of the User Interface components and their grouping into contexts in the CUSP software aligns with these principles.

The context of **task selection** is a prime example of design efficiency. The main taskbar at the top of the screen allows users to easily switch between tasks such as *Trends*, *Events*, *Prediction*, *Actuation*, *Optimisation*, and *Composition*. This design reduces the number of steps required to switch tasks, thus increasing efficiency. Users do not need to navigate through complex menus or interfaces to access different functions, making the process quick and straightforward.

The **Entity Selection Context** also exemplifies efficiency and effectiveness. Users can select entities from a directory on the left taskbar or by clicking on a 3D model within the viewer surface. This dual-method approach caters to different user preferences and work styles, thus increasing efficiency. The Directory is organised based on semantic ontology, making it easier for users to find and select entities. This design speeds up the entity selection process and reduces errors, improving effectiveness.

The **Layer Manipulation Context** is another example of efficiency in design. The Layers tab allows users to toggle the display of low- and high-resolution models. This feature gives users control over the level of detail they want to view, allowing them to quickly switch between detailed and overview perspectives. This flexibility can increase efficiency, particularly in tasks that require a broad overview or detailed inspection.

The **Graph Generation Context** enhances both efficiency and effectiveness. Users can generate graphs based on various parameters using the Graphs component. The interactive nature of the graphs, provided by the ability to maximise, zoom, adjust focus, and reset the view, allows users to manipulate the data representation to suit their needs. This makes not only the task of data analysis more efficient but also more effective, as users can tailor the graphs to reveal the most relevant and insightful information.

The design of the User Interface components and their grouping into contexts in the CUSP software demonstrates a solid commitment to efficiency and effectiveness. By reducing the number of steps required to perform tasks, offering flexible and powerful tools, and minimising the potential for errors, the software enables users to work more productively and achieve their goals more accurately.

5.4.4 Enjoyability

Hassenzahl, Mahlke, and Karapanos emphasise creating an enjoyable user experience. The Graph Generation Context in the CUSP software contributes to this by allowing users to interact with the graphs dynamically. Users can maximise the graph, zoom in on specific areas, adjust the focus on the timeline, and reset the view of the graph. These interactive features make the experience of using the software more engaging and pleasant.

Enjoyability is a crucial aspect of User Experience (UX) design, as highlighted by Hassenzahl, Mahlke, and Karapanos. An enjoyable user interface is not only functional and efficient, but also engaging and satisfying to use. The design of the User Interface components and their grouping into contexts in the CUSP software aligns with this principle.

The **Graph Generation Context** is a prime example of how enjoyability is incorporated into design. The Graphs component allows users to generate and interact with graphs dynamically and engagingly. Users can maximise the graph, zoom in on specific areas, adjust the focus on the timeline, and reset the view of the graph. These interactive features make the experience of using the software more engaging and pleasant. The ability to interact with the data hands-on can lead to a deeper understanding and appreciation of the data, enhancing the user's satisfaction.

The **Building Visualisation Context** also contributes to the enjoyability of the software. The 3D viewer surface supports panning, zooming, and orbiting, allowing users to interact with the buildings naturally and intuitively. This interactive 3D visualisation provides a practical way to view and navigate the building models and creates a visually engaging and immersive experience that enhances the enjoyment of using the software.

The Layer Manipulation Context adds another layer of enjoyability. The Layers tab allows users to toggle the display of low- and high-resolution models. This feature gives users control over the level of detail they want to view, allowing them to switch between a broad overview and a detailed view. This flexibility can make exploring models more enjoyable, as users can tailor the view to their preferences and interests.

The **Prediction Scenario Context** enhances the ease of use of the software by allowing users to generate prediction scenarios and view the settings of the selected scenario. This feature gives users a sense of control and participation, as they can actively participate in the prediction process and see the impact of their decisions. The ability to visualise the corresponding carbon emissions for a scenario adds another layer of engagement and satisfaction, as users can see the potential environmental impact of their decisions.

Therefore, the design of the User Interface components and their grouping into contexts in the CUSP software demonstrates a solid commitment to enjoyability. By providing engaging, interactive, and visually appealing features, the software creates a functional and efficient user experience, but also enjoyable and satisfying.

5.4.5 Predictability and Consistency

Norman's principles also emphasise the importance of predictability and consistency in design. The IFC File Import Context and the Asset Upload Context in the CUSP software follow these principles by providing consistent and predictable interfaces for uploading files. The forms used in these contexts prompt the user to select some properties, the type of file, and the entity to attach the file, providing a consistent user experience.

Predictability and consistency are fundamental principles of good usability, as highlighted by Don Norman. A predictable interface allows users to understand what will happen before they act, while a consistent interface ensures that similar elements function similarly throughout the system. The design of the User Interface components and their grouping into contexts in the CUSP software aligns with these principles.

The **Task Selection Context** is a prime example of predictability in design. The main taskbar at the top of the screen allows users to easily switch between tasks such as *Trends*, *Events*, *Prediction*, *Actuation*, *Optimisation*, and *Composition*. The consistent placement and functionality of the taskbar on different screens make it predictable to users, reducing cognitive load, and enhancing usability.

The **Entity Selection Context** also exemplifies predictability and consistency. Users can select entities from a directory on the left taskbar or by clicking on a 3D model within the viewer surface. This dual-method approach is applied consistently across the software, making the entity selection process predictable and easy to learn.

The **Layer Manipulation Context** is another example of predictability in design. The Layers tab allows users to toggle the display of low- and high-resolution models. This feature is consistently available, allowing users to predictably control the level of detail they want to view, regardless of where they are in the software.

The IFC File Import Context and the Asset Upload Context demonstrate consistency in design. Both contexts provide forms that guide the user through uploading a file. The consistent design of these forms, including the prompts to select some properties, the type of file, and the entity to attach the file, provides a predictable user experience, reducing the likelihood of errors, and improving usability.

In summary, the design of the User Interface components and their grouping into contexts in the CUSP software demonstrate a solid commitment to predictability and consistency. CUSP improves usability and user satisfaction by ensuring that similar elements function in the same way throughout the system and that users can predict the outcome of their actions.

5.4.6 Feedback and Visibility

Another essential principle of good usability is providing feedback and visibility of the status of the system, which are crucial principles of good usability, as highlighted by Don Norman. Feedback ensures that users are informed about the results of their actions, while visibility of system status keeps users informed about what is happening within the system. The design of the User Interface components and their grouping into contexts in the CUSP software aligns with these principles.

The **Building Visualisation Context** provides clear feedback through the Building Load Progress indicator. As users load a building model, they can see the progress of the loading process, providing immediate feedback on their actions. This feedback keeps users informed about their request's status and gives them a sense of control and assurance that the system is responding to their actions.

The **Graph Generation Context** also provides feedback and visibility. As users interact with the graphs, such as zooming in on specific areas or adjusting the focus on the timeline, they can see the immediate impact of their actions on the graph. This real-time feedback enhances the interactivity of the graphs and makes the data exploration process more engaging and informative.

The **Prediction Scenario Context** enhances the visibility of the system status by displaying the settings of the selected prediction scenario. This feature provides users with a clear understanding of the current configuration of the prediction scenario, keeping them informed of the status of the system. The visibility of the scenario settings also helps users understand the potential impact of their decisions, enhancing their decision-making process.

The **Asset Uploader Context** provides feedback throughout the file upload process. The upload process begins after users select a file from their local system and click "Upload". Users can see the progress of the upload, providing immediate feedback on their actions. Once the upload is complete, the file appears as an attached asset to the selected entity, providing further feedback that the upload was successful.

As such, the design of the User Interface components and their grouping into contexts in the CUSP software demonstrate a solid commitment to feedback and visibility of system status. Software improves usability and user satisfaction by informing users about the results of their actions and the status of the system.

5.4.7 Error Prevention

Norman's principles also emphasise error prevention. The software achieves this through forms in the IFC File Import Context and the Asset Upload Context, which guide the user through uploading a file, reducing the likelihood of errors.

Error prevention is a crucial principle of good usability, as highlighted by Don Norman. A well-designed interface should help users recover from errors and prevent errors from occurring in the first place. The design of the User Interface components and their grouping into contexts in the CUSP software aligns with this principle.

The **Entity Selection Context** is a prime example of error prevention in design. Users can select entities from a directory on the left taskbar or by clicking on a 3D model within the viewer surface. This dual-method approach reduces the likelihood of errors in entity selection. For instance, if users struggle to select an entity by clicking on the 3D model due to its size or position, they can easily select the same entity from the Directory. This redundancy in selection methods helps to prevent errors and improve usability.

The **IFC File Import Context** and the **Asset Upload Context** also demonstrate error prevention in the design. Both contexts provide forms that guide the user through uploading a file. The forms ask the user to select some properties, the type of file, and the entity to attach the file. By guiding the user step by step through the process, these forms help prevent errors, such as selecting the wrong file type or attaching the file to the wrong entity. This guide makes the upload process more accessible and reduces the likelihood of errors, enhancing the overall user experience.

The **Layer Manipulation Context** is another example of error prevention. The Layers tab allows users to toggle the display of low- and high-resolution models. By providing a clear and straightforward way to switch between different levels of detail, this feature helps prevent errors that could occur if users had to manually load different models or navigate through complex menus to change the level of detail.

The **Prediction Scenario Context** enhances error prevention by displaying the settings for the selected prediction scenario. This visibility helps users understand the current configuration of the scenario and prevents errors that could occur from misunderstanding the scenario settings. For instance, users can see whether the scenario includes visualising corresponding carbon emissions, preventing them from mistakenly assuming that this feature is included or excluded.

In conclusion, the design of the User Interface components and their grouping into contexts in the CUSP software is a thoughtful application of User Experience and sound usability principles. The design is user-centred, efficient, pleasant, predictable, and consistent and provides feedback and visibility of the status of the system, contributing to an overall positive user experience.

5.4.8 Code conversion to VueJS framework

A subsequent step that we took, as part of one of the iterations of the "continuous deployment and improvement" stage, the researcher migrated the front-end framework to VueJS. VueJS is an open source JavaScript framework that has recently gained substantial popularity, having been integrated as part of a suite of creative web applications, and products such as Airbnb, Github, and Instagram have already been developed [372].

The contemporary world of web development demands a dynamic and fluid user interface, highly scalable solutions, and accessible and efficient codebases. The Vue.js framework, a modern JavaScript library for building web interfaces, offers all of this through its innovative design and functional architecture. This subsection argues that the Vue.js framework adheres to principles and best practices of good user experience (UX), usability, and scalability in web applications, substantiating the claim with academic and practical perspectives.

The Vue.js framework promotes good UX by ensuring seamless, fluid, and interactive web application designs. Its component-based structure allows developers to build reusable and encapsulated parts that maintain their state, leading to a more modular and manageable code base [235]. This approach is consistent with Nielsen's principles of modularity and flexibility in user interface design [260]. Furthermore, Vue.js includes transition components, which can automatically apply transition effects when elements are inserted, updated, or removed from the DOM [235]. These effects contribute significantly to visual feedback, satisfying the UX principle of immediate response [265].

Vue.js prioritises usability through simplicity and accessibility. Its learning curve is much gentler than other contemporary frameworks, such as Angular or React, making it more accessible to developers of different skill levels [299]. This accessibility improves the learnability of the software, a critical aspect of the usability of the system [261]. Furthermore, Vue.js provides detailed documentation, making it easier for developers to find solutions to problems and learn new features [235]. High-quality documentation improves the usability of a software by supporting user needs [1].

Finally, the Vue.js framework is built for scalability. Due to its progressive nature, it can handle both small-scale projects and larger, more complex applications [235]. This attribute aligns with the best practices of scalable software design outlined by O'Reilly and Binns [273], including modularity and the ability to handle increasing demands with increased resources. Additionally, since Vue.js uses a virtual DOM and has a reactive data binding model, it allows efficient rendering and updating of views, contributing to performance optimisation and scalability [78].

In conclusion, the Vue.js framework aligns with the principles and best practices of UX, usability, and scalability, affirming its position as a robust and reliable tool for web application development. The framework's component-based structure and transition components promote good UX, while its learning accessibility and comprehensive documentation foster usability. Its progressive nature, virtual DOM, and data-binding model support scalability. Therefore, Vue.js offers a multifaceted and adaptable solution for modern web development.

This chapter used semantics to explore the critical aspect of providing context-informed digital twins, as it aimed to answer the second research question: *How can we derive context from static artefacts (buildings) and dynamic artefacts using a Digital Twin?*. Focusing on the Computational Urban Sustainability Platform (CUSP), this chapter highlights the software architecture comprising IoT sensors, a semantic middleware, data science services, and the User Interface. Integrating semantic structures and ontologies enables context derivation from static artefacts like buildings and dynamic artefacts facilitated by digital twins. With the intuitive result of the Web user interface and visually annotated data structures, the CUSP platform allows users to interact effortlessly with the digital twin, extract valuable insights, and make informed decisions. Using semantics, CUSP aims to improve user experience, foster stakeholder collaboration, and optimise communication in BIM and CIM Web applications, paving the way for a robust context-aware user experience in urban sustainability decision support tools.

5.5 Variation of User Experience aspects

The Computational Urban Sustainability Platform (CUSP), a cornerstone of the research presented in Section 5, serves as a testament to the integration of semantic structures and ontologies to derive context from static and dynamic artefacts. The following table encapsulates the multifaceted aspects of user experience within the CUSP user interface, offering a comprehensive summary that correlates UI contexts with the associated user interactions and benefits. This synthesis is instrumental in illustrating the tangible impacts of semantic enhancement on user experience and interaction within digital twins.

The detailed exposition of user experience aspects within CUSP's UI contexts underscores the intricate relationship between semantic integration and user interaction. It is evident that the iterative development and participatory evaluation of CUSP have culminated in a platform that not only facilitates context-aware data interpretation and decision making, but also fosters a user experience that is intuitive and adaptable to varying user needs and scenarios. As we transition to the subsection on the evolution of the bespoke UX model, it is important to recognise how these detailed UI components and interactions contribute to the larger narrative of enhancing user engagement and satisfaction through semantic empowerment and digital twinning. The forthcoming analysis will further delve into how these methodologies and tools translate into a refined user experience, reinforcing the digital twin's role as a bridge between the digital and physical realms.

Table 4: User Experience Aspects in CUSP UI Contexts

UI Context	Primary Components	User Actions	User Benefits	Variations for Different Scenarios
Authentication	Authentication screen	Log into the system	Secure and simplified user experience	Different authentication methods for varied security levels
Building and World Visualisation	3D viewer, World Visualiser	View 3D models and their global position	Enhanced spatial understanding	Customisable viewing angles and perspectives for diverse user roles
Task Selection	Main taskbar	Switch between different tasks	Increased efficiency and reduced cognitive load	Taskbar customisation based on user preferences or role
Entity Selection	Directory, 3D model selection	Navigate hierarchy, select entities	Intuitive data structure interaction	Adaptive directory structure for complex ontologies
Layer Manipulation	Layers tab	Toggle display of model details	Control over detail level, improved clarity	Dynamic layer management for large-scale models
Information Viewing	Information button	Access detailed building information	Informed decision-making	Context-sensitive information display
Graph Generation	Graphs component, data stream highlight	Generate and interact with graphs	Insightful data visualisation	Configurable graphs for specialised data analysis tasks
Simulation Creation	Simulation tools	Run Energy Plus simulations	Tools to support specific simulation tasks	Scalable simulation features for varying complexity
IFC File Import	IFC uploader form	Import IFC files into ontology	Integration of external building data	Import wizards for different data formats
Asset Uploader	Asset Uploader form	Upload assets to entities	Streamlined upload process	Multi-file and bulk upload capabilities
Prediction Scenario	Scenario settings display, cost-benefit ratio, carbon emissions checkbox	Generate and evaluate prediction scenarios	Strategic decision support	Scenario templates for common use cases

In the continuum of this thesis, following the discourse on delivering context-informed Digital Twins through semantics, a table summarising the User Experience Aspects in the CUSP UI contexts is presented. This table provides a granular view of the interaction between users and the system, detailing the primary components, user actions, benefits, and variations tailored for different scenarios. The inclusion of this table is pivotal as it elucidates the practical applications of theoretical constructs discussed in the previous sections and sets the stage for the forthcoming examination of the evolution of the bespoke UX model.

5.6 Evolution of the bespoke UX model

The evolution of the bespoke UX model encapsulated enhancing user interaction with semantics, a critical step in the context-informed Digital Twins and semantic web applications domain. Through the integration of semantic technologies, users' interactions with BIM web applications underwent a significant transformation, culminating in more intuitive, efficient, and meaningful experiences. This process of enhancement involved several vital aspects:

Semantic integration in data interpretation emerged as the cornerstone of the UX model, using semantic technologies to present data and interpret them within their context. This meant that raw data from IoT sensors in a building, for example, was not shown as mere numbers but translated into valuable insights related to energy efficiency, space use, or maintenance needs, thus giving users a meaningful understanding of the information. Additionally, the model shone by interpreting dynamic data and adjusting outputs in real time, which proved indispensable in the ever-changing landscapes of complex BIM projects.

User-centric semantic queries were a leap forward, with the introduction of natural language processing (NLP) capabilities that allowed users to engage with the system using natural conversational language, a boon for usability, particularly for those without expertise in semantic languages. Intelligent search functionalities were refined to understand the intent behind user queries, delivering pertinent results even when questions were imprecisely articulated.

The visualisation of semantic data was also pivotal, with data encapsulated in intuitive dashboards that made the understanding of complex information straightforward; recognisable patterns, trends, and anomalies could be discerned at a glance. Moreover, the model included interactive elements, such as clickable components and real-time updates, transforming data exploration from a passive, laborious task into an active, captivating experience.

Personalisation and recommendation further bespoke the UX model through User Profile Learning – an adaptive feature that refined the system's understanding, tailoring its behaviour, and advice to match individual user preferences and work habits. Through proactive recommendations, the model went further to offer data and foresighted suggestions that could identify optimisations and preemptively flag potential problems.

Semantic collaboration tools were introduced, enabling users to collaborate on models and documents with semantic tags, bolstering collective understanding and ensuring team alignment.

Including versioning and change tracking was instrumental for collaborative projects, providing clarity on the evolution of BIM models and the specific contributors to changes.

Integration with external semantic resources linked the model with abundant external semantic data and ontologies, enriching the data set, and ensuring that the furnished information was comprehensive and current. The model also aligned with industry standards, ensuring regulatory compliance and adherence to best practices.

Enhancing user interaction with semantics was transformative for the UX model, transitioning it into a powerful instrument that transcended mere data representation, and providing users with context-rich, insightful, and actionable guidance. The model propelled users towards a more profound and more instinctive synergy with BIM web applications by harnessing semantic technologies, significantly contributing to well-informed decision-making and adept project management within the built environment.

6 Scaling up semantics at the city level with citizens as active sensors

This chapter delves deeply into the progressive concept of transforming urban landscapes into dynamic, intelligent ecosystems in which citizens actively participate as sensors. Building upon the theoretical foundation established in previous chapters, we apply the principles of semantic web technology within a broader scope, specifically within urban environments. This marks a significant stride towards the broader goal of enriching user experiences across multiple applications in the built environment.

At the heart of the discussion lies the captivating notion of citizens as active sensors, individuals who observe, interpret, and interact with their surroundings, generating valuable semantic data. These active sensors play a pivotal role in shaping the dynamic context of a city, providing tangible and intangible data. Tangible data encompasses physical metrics such as traffic patterns, weather conditions, and architectural structures. In contrast, intangible data encompass more abstract elements, including social interactions, cultural nuances, and emotional responses to spatial configurations.

To address the central research question (*how can we harness this multifaceted dynamic context to enhance user experiences in various applications within the built environment?*) the researcher explores combining perspectives from the realms of humanities and technology. Whether it is the excitement of a spectator in a stadium, the strategic decisions of a city planner, or the overarching vision of a city manager, the convergence of semantics and human experiences expands the possibilities for gaining insights into the surroundings.

Through a series of case studies, conceptual models, and empirical data, this chapter presents a comprehensive understanding of how citizens, acting as active sensors, contribute to the semantic layer of a city. By considering various scales, from specific building neighbourhoods to the entire city, the researcher leverages the ability to interpret and respond to these data, potentially leading to more responsive, inclusive, and adaptive cities. This represents a significant step towards creating real “smart” urban environments.

This chapter is divided into several key subsections, each addressing a subresearch question that contributes to the researcher’s study.

The first subsection, ”Harnessing Natural Language Processing (NLP) for Semantic Interpretation of Social Media Data,” explores how NLP techniques can decipher the vast stream of social media data in a smart-city context. This subsection investigates the intersection of social media and smart cities, the role of citizens in data generation, and the application of advanced NLP techniques to process and understand these data.

The subsequent subsection “Exploring Patterns and Trends in Citizen Satisfaction within a Smart City Context” shifts the focus to uncovering general patterns and trends in Citizen Satisfaction within a smart city context. Through a comprehensive analysis of quantitative and qualitative data, this section sheds light on discernible patterns and trends over time and across different urban contexts, offering insight into citizen satisfaction.

Finally, the third subsection, “Interlinks between Trends in Citizen Satisfaction and Concurrent Environmental Factors and Events” examines the emerging relationships between the aforementioned trends in Citizen Satisfaction and Concurrent Environmental Factors and Events. The researcher extensively investigates how environmental variables and significant city events interact with and influence citizen satisfaction trends, providing an integrated view of how semantics and context converge to shape user experiences within a smart city.

By addressing these subresearch questions, the researcher aims to paint a comprehensive picture of how semantics can be applied at the city level, with citizens acting as active sensors.

6.1 Harnessing Natural Language Processing (NLP) for Semantic Interpretation of Social Media Data

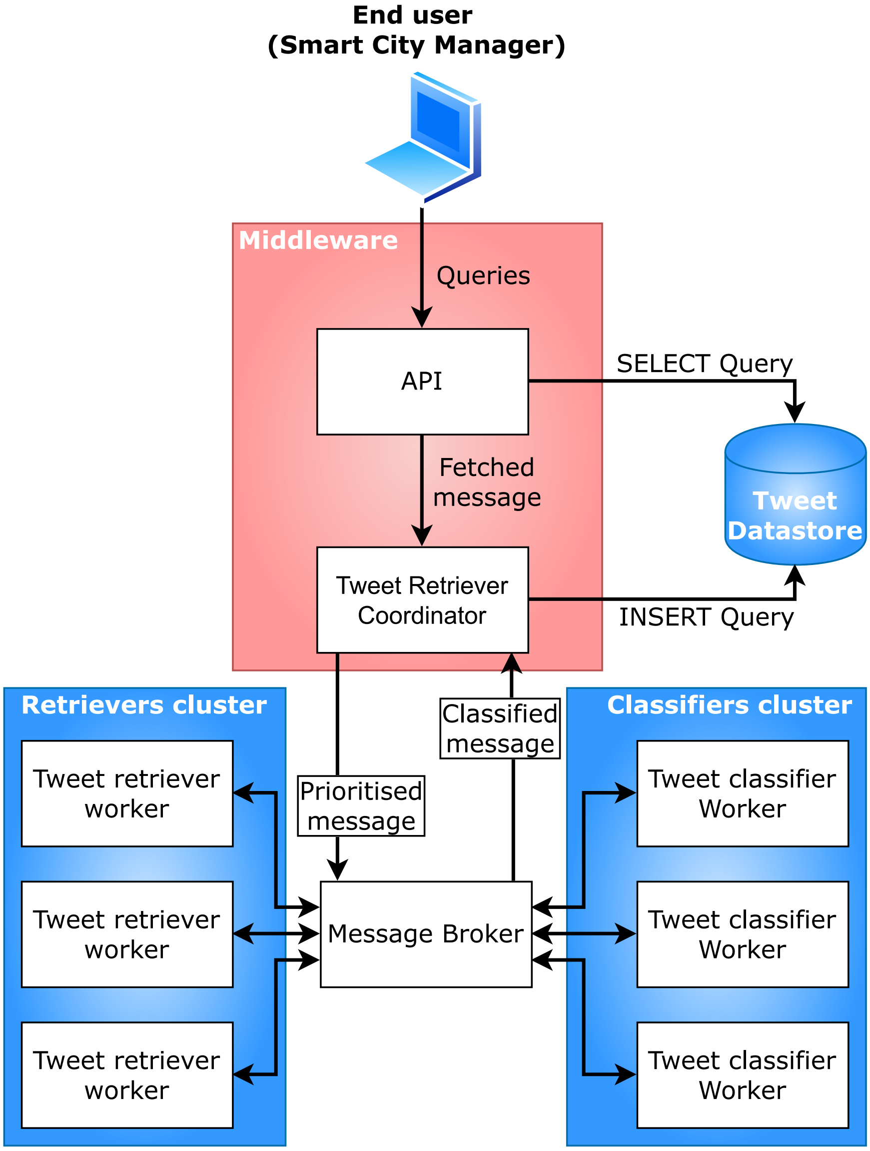
Within the first subsection, the researcher delves into the realm of Natural Language Processing (NLP) and its application to interpret the deluge of social media data within a smart city context. The ubiquitous presence and extensive use of social media platforms have created an ever-flowing data stream that can offer invaluable insights into the urban landscape when mined and analysed correctly. The researcher explores how NLP, a subfield of artificial intelligence that deals with the interaction between computers and humans using natural language, can help extract, understand, and use these insights. This inquiry provides a foundation for using NLP techniques to transform raw, unstructured data from citizens into meaningful and actionable semantic information.

6.1.1 Data pipeline and application architecture

The proposed application architecture aims to operationalise the methodology outlined in Chapter 4, which aims to address the research question and the subresearch questions outlined above.

Figure 34 presents a high-level architectural overview of the technical setup. We designed a redundant data pipeline to speed up retrieval, spreading the workload over multiple machines.

Figure 34: Data pipeline within the three modules of the Application (*Middle-ware*, *Retrievers cluster*, *Classifiers cluster*). The *Tweet Retriever Coordinator* splits the messaging interval into smaller chunks and publishes them to the *Message Broker*.



To abstract the underlying OS (Operating System) for uniformity and resource compartmentalisation, a virtual network of *Docker* containers [81] was deployed using *Debian*-based Linux image environments. Users can query data through the *API (Application Programming Interface)*, specifying the date interval, area, and other parameters. This architectural implementation, coupled with a *Message Broker* conducting parallel processing aided by the RabbitMQ library [170], allowed us to handle the collection and classification of tweets in (nearly) real time.

We executed ten parallel classifier instances as part of a High-Performance Computing (HPC) cluster of machines. The delegation of processing tasks takes into account: (a) the length of the message needing to be processed, (b) the semantic complexity of the keywords used as part of the tweet message, (c) the presence of polluting factors, such as stop words, punctuation, orthography, quotes from movies, music lyrics, para-verbal written language (emoticons/smileys).

The *Tweet Retriever Coordinator* is a simple automated Python script listening to the queue of published intervals while simultaneously connecting to Twitter and downloading tweets matching specific criteria (most commonly, time intervals, keywords and geo-coordinates). This process is scheduled to publish all intervals for the required historical period on to the *Message Broker*.

After retrieval, each tweet is published as an individual message to the *Message Broker* to be processed. The *Message Broker* is the central point of the application, allowing for quick synchronisation between processes and message buffering while avoiding workload duplication. Workers would request the processing of new messages at their own pace in a non-symmetrical usage model, redistributing their load into multiple segments running on different machines when their I/O (Input/Output) usage passes a certain threshold. After an interval, the resulting tweets would be queued as individual messages to be processed.

This language model is trained to predict the next word based on the previous words taken as input. During this phase, the neural network understands the English language and vocabulary. The *category* model is then trained on the data set, according to the methods described in Section 6.1.2. Once the language models are trained on the dataset, we train the *Classifiers*, comprising linear layers placed at the end of the encoder of the language model. Each classifier instance loads the two models (*category* and *emotional predilection*), then requests messages to be processed and determines the label and sentiment for each of them.

To store the models and preload them, we used the *Pickle* serialisation library [117]. Upon receiving a message, the text is preprocessed, classified, and finally published back to the *Message Broker*. To avoid encumbering the *Classifier*, we stored the curated tweets in a separate database, together with the messages retrieved from the *Message Broker*. The asynchronous mode operandi allowed the *Classifier* to work at total capacity despite the I/O (Input / Output) load.

6.1.2 Dataset construction and processing

As Kalinin et al. [178] mentioned, smart cities are about preventing risks in addition to analysing the quality of services and procedures that differ from city to city. Therefore, we carefully selected a classification with as broad an applicability spectrum as possible. This approach enables us to demonstrate the viability of the technical solution beyond the confines of the case study illustrated in Section 3.5.1.2 for data filtering and classification. The *Classifier* inherits the same seven event types as in this *Taxonomy*.

6.1.2.1 Processing models and pre-emptive testing

Before embarking on tweet collection and processing, we considered a baseline model for event detection. We preemptively evaluated several language classification models identified in the literature review. For baselines, tweets are converted into a matrix of token counts before considering the following three types of classifier (language models used for classifying texts):

- *MNB (Multinomial Naive Bayes Classifier)*, computing the probabilities of belonging to a class as a function of the occurrence of different words (of high popularity due to its simplicity).

- *CNB (Complement Naive Bayes Classifier)*, utilising the same principle behind MNB while correcting its assumptions and making it suitable for imbalanced data (better suited for conversational rather than informative tweets).

- *RF (Random Forest Classifier)*, providing accurate baselines for regression and classification tasks.

We tested the *Classifiers* using the *Scikit Learn* machine learning library [197], with hyperparameters defined as follows: $alpha = 0.5$ for *MNB*; $alpha = 1.5$ for *CNB*; $n\_estimators = 50$ for *RF*. The primary comparison metric is the accuracy, measured more than *five times.* As expected, the CNB classifier outperforms the MNB. Models incorporating a preprocessing stage produced more accurate predictions than their counterparts processing raw data. Some complex hashtags are altered and favour the model without preprocessing because it can label tweets containing the same hashtag. However, we observed that this gap narrowed as we collected additional data, reducing the bias induced by the classifier to the maximum extent possible.

As a classification model for the data mining and classification stages, we use UMLFiT (Universal Language Model Fine-Tuning), a technique created by devised by Howard and Ruder [162], which pre-trains an LM (Language Model) on a sizeable general-domain corpus and then uses novel techniques to fine-tune it for the target task. The method is universal in the sense that it meets the following practical requirements: (i) it applies to tasks with varying document sizes, numbers, and types of labels; (ii) it is built on a single architecture and training process; (iii) no custom feature engineering or preprocessing is required; and (iv) no additional in-domain documents or labels are required. ULMFiT comprises the following steps:

1. *General-Domain LM Pretraining*: the LM is pre-trained on a large general-domain corpus to forecast the following word in a sequence (with a particular degree of certainty). At this stage, the model acquires knowledge of general linguistic characteristics, such as the typical sentence structure of “subject-verb-object” of an English sentence.

2. *Target Task LM Fine-Tuning*: the LM is fine-tuned on the data of the target task, acquiring task-specific features of the language elements (in the case of Twitter messages), such as the existence of handles, the usage of slang, abbreviated words, and emojis.

3. *Target Task Classifier*: as a third step, the pre-trained LM is expanded by two linear blocks so that the final output is a probability distribution over the sentiment labels (i.e. *positive* or *negative*), and a label from the risk taxonomy, respectively.

For the pre-training steps of ULMFiT, we propose the innovative language model AWD-LSTM Merity et al. [241], which encompasses a standard LSTM (Long Short-Term Memory language model) with various tuned dropout hyperparameters (which involve no additional complexity layers, such as attention or shortcut connections). ULMFiT enables the adaptation of a pre-trained model, fine-tuning the neural network layers to tailor it to the research objectives. ULMFiT is integrated with the dataset through the *Python*-based *fast.ai* library [161], which facilitates freezing / unfreezing of layers, offering a fine-tuned approach to customising the model.

Table 5: Accuracy of the previously utilised classification models vs the combination (*AWD-LSTM* and *ULMFiT*)

Model type	*Without tweet pre-processing*	*With tweet pre-processing*
MNB	78.2% (± 3.4%)	76.6% (± 4.6%)
CNB	80.2% (± 2.7%)	79.6% (± 3.6%)
RF	85.0% (± 5.4%)	83.2% (± 6.6%)
AWD-LSTM & ULMFiT	**88.5% (± 3.2%)**	**88.4% (± 2.0%)**

As we can observe in Table 5, this combination of techniques has exhibited a higher precision than its counterparts according to the measured parameters. Although acknowledging the susceptibility to overfitting of this approach, the researcher has mitigated this by employing a *stratified K-fold cross-validation.* More information on the technical setup that underpins the association of these two NLP techniques can be found in a recent article by Briskilal and Subalalitha [62].

6.1.2.2 Dataset scope, sourcing and filtering

Our target dataset consisted of messages broadcast by people residing in or transitioning through the City of Cardiff, United Kingdom. For the purpose of the manuscript, we only analysed tweets written in English.

Table 6: The filtering Cardiff City bounding box coordinates (TLP = Top-Left Point, BRP = Bottom-Right Point)

TLP Longitude	*TLP Latitude*	*BRP Longitude*	*BRP Latitude*
−3.325587	51.444454	−3.067674	51.554148

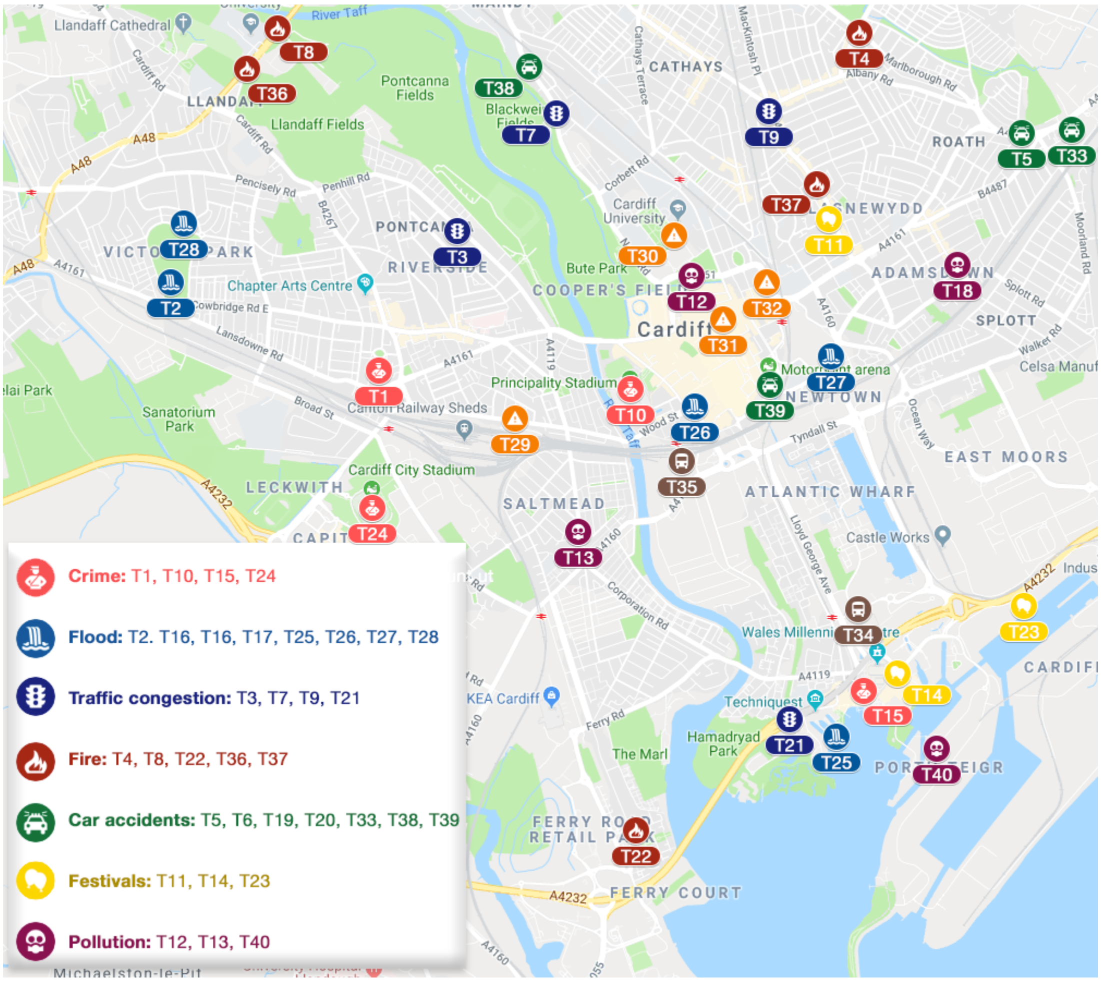
A geofencing boundary box (Table 6) was defined to filter the live data stream, so only Cardiff tweets were collected.

The *twint* Python library [381] was used to retrieve the tweets, enabling timely collection without utilising the Twitter API (Application Programming Interface), which could have been subject to call limitations. For data anonymisation purposes, once a message is received, only a generic unique *id* of the tweet, *date* and *text* was stored in the database to avoid any links that could be established with the identities of actual persons.

Although a significant proportion of the collected tweets were not geolocated (because users did not have the "location sharing" function enabled on their devices), we determined the location of the users through other means wherever possible. Another mini neural network was trained to predict originating neighbourhoods or landmarks based on the *biography* of the user and additional information shared within their profile (connections, hashtags, and other references to locations in their tweets).

An initial test data subset generated the citizen-accessible interactive map illustrated in Figure 35. Following the initial promising results, we proceeded to deploy the production environment for tweet collection. 1.6 million tweets over a nine-month period in 2020.

Figure 35: The interactive map of events in the city of Cardiff derived from the initial results produced by the classifiers



6.1.2.3 Dataset pre-processing

Before feeding the contents of the tweets into the model, we pre-processed them to make them ready for parsing by the tokeniser. This process represents an additional safeguard to ensure that as much information is retained from the original content as possible while also avoiding classification bias occurring within the model (for instance, associating an event to a specific date). Tweets were filtered using keywords and hashtags, adjusting the generality of filtering rules to prevent inducing bias in the training process of the *Classifier*. Tweets are saved within the data store alongside their semantic selection criteria (the keywords and/or hashtags used for identification). To account for lexical diversity, similar tweets were further filtered using the *cosine similarity* approach [382].

For the text of each tweet: (i) encoding errors are repaired; (ii) retweets, hyperlinks, emails, phone numbers, “hashtags”, “cashtags”, signs, date, time, smileys and emojis are replaced by unique tokens; (iii) text written in the “CamelCase” syntax is split, and unique tokens are added at the start of the corresponding expressions; (iv) unnecessary white spaces are removed. The researcher increases the quality of this process by using carefully crafted regular expressions, which minimise the number of false positives. Before inputting the text into the model, the final step is to assign tokens using the *fast.ai* tokeniser (which also provides access to the ULMFiT language model).

6.2 Exploring Patterns and Trends in Citizen Satisfaction within a Smart City Context

This subsection focuses on the patterns and trends that can be observed in citizen satisfaction in the context of a smart city. With citizens acting as active sensors, their feedback, expressed through various channels and formats, becomes essential for understanding their experiences and satisfaction levels. The researcher conducts a comprehensive quantitative and qualitative data analysis that represents diverse individual experiences, collective sentiments, and complex emotional landscapes. By identifying and interpreting these patterns and trends, the researcher aims to provide a nuanced understanding of citizen satisfaction, thus contributing towards a more informed approach to improving user experiences in smart cities.

6.2.1 Multiple Regression Analysis applied to the trends identified as part of the scenarios

Being a supervised learning task by itself, we consider MRA a highly suitable and accessible statistical instrument to validate the preliminary correlations determined by the supervised learning NLP techniques. Given the predominantly linear nature of the relationships between the variables forming the researcher’s scenarios, MRA was considered to strike a favourable compromise between complexity and accuracy. The results are easier to interpret, in contrast with the output of other algorithms. The stages of MRA validation were tailored to the research objectives and independently applied to each scenario.

6.2.1.1 General regression statistics and ANOVA

The following general regression statistics are applied to all scenarios: *Multiple Determination Ratio R*, R^2, *adjusted* $R^2 = (\sqrt{\frac{ESS}{TSS}})$, and *Standard Error (SE)*. *Multiple R* indicates the intensity and nature of the connection between the variables. The *Adjusted* R^2 indicates the percentage of variation of the independent variable that the simultaneous variation of the independent variables can explain. Unlike the R^2, the *Adjusted* R^2 considers the degrees of freedom and the number of parameters included in the model.

Standard Error (SE) is the approximate standard deviation of a statistical sample population, which measures the precision of a representation of a sample distribution of a sample population. A "sample mean" deviates from the actual mean of a population sample, and this deviation is the standard error of the mean. The *SE* can also be explained as the difference between the expected and actual values of the variables. All scenarios exhibited a relatively small *SE*, which is an indication that the sample mean is an accurate reflection of the actual population mean.

For each scenario, we analyse the significance of the model using the *Fisher test* applied to the *ANOVA (Analysis of Variance) table*. ANOVA is a statistical analysis tool used to test the degree of differences between two or more groups that are part of an experiment. The results of the ANOVA test are displayed in a tabular form known as an ANOVA table, which displays the statistics used to test hypotheses about population means (testing them with intragroup and intergroup variants). For each scenario, we formulate the following two hypotheses: (i) $H_0 : \beta_0 = \beta_1 = \beta_2 = 0$ (the model is not statistically significant) and (ii) $H_1 : \beta_0 \neq 0$ or $\beta_1 \neq 0$ or $\beta_2 \neq 0$ (the model significantly explains the connection between the variables). Based on the value of the *Fisher test (f-test)*, we decide whether to reject hypothesis H_0 and what risk we apply. If we reject H_0, we conclude that the model is statistically significant (for 5% risk). Then we broadly determine the type of dependence between the variables.

6.2.1.2 Analysis of regression coefficients and variability for significant parameters

The second step for analysing each scenario is applying the *Student's* $t - test$ on the regression coefficients, considering the estimators obtained using the *Least Squares* method and their distribution law. Firstly, we write the equation of the multiple linear regression model using the following formulas: $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$ and $Y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3$. In both formulas x_1, x_2 and x_3 are the values of the *independent variables*, Y is the value of the *dependent variable*. In the first formula, β_0, β_1, β_2, and β_3 are the *hypotheses variables* associated with the significance of each parameter. In the second formula, b_0 is the *coefficient of the dependent variable*, and b_1, b_2, b_3 are the *coefficients of the independent variables*.

For each parameter β above, we follow the classical steps of the MRA statistical evaluation:

1. formulating the two hypotheses ($H_0 : \beta\ = 0$ and $H_1 : \beta\ \neq 0$)

2. selecting a significance threshold $\alpha = 0.05$

3. choosing the statistical test t and an associated theoretical value of the statistic $t_{\frac{\alpha}{2},n-k}$

4. calculating the test value for the associated coefficients: $\frac{b_1}{s_{\hat{\beta 1}}}$

5. for each parameter in the linear regression model equation, we apply the following decision rule: if $\sigma < \alpha$, we reject H_0, and we conclude that the parameter β is statistically significant (for 5% risk)

6. for each parameter identified as "significant", we showcase the correlation between its variability and the variability of the dependent variable, quantifying the variations that occur on average

6.2.1.3 Residuals

To further determine the statistical significance of the regression models corresponding to each scenario and further verify their accuracy and relevance, we analyse the following metrics and plots where appropriate: (i) plots of residuals vs fitted values, (ii) case order plots of leverage, (iii) normal probability plot of residuals, (iv) histogram plot of residuals (v) residuals vs observation order.

The scenarios analysed within this section use events previously detected following the data pipeline defined in Section 6.1.1 and leverage the algorithms and methods defined in Section 6.1.2. The motivation and reasoning behind the selection of scenarios have been described in Section 3.5.1.

At the beginning of the analysis of each scenario, we outline the general trends in the number of occurrences of risk-based events on a daily time frame, present within the evaluated period in the year 2020. The researcher first inferred a subjective correlation between the dependent and independent variables based on these general trends. In the second part of each scenario analysis, the researcher statistically validated the previously inferred subjective correlations by applying MRA statistical models to the classified datasets.

As we can observe from the graphical trend representations outlining the occurrences in the sections below, Scenarios 1-5 consistently featured an intermittent reduction in activity between August and September. This pattern could be explained by uncertainties and restrictions related to COVID, cooccurring with the start of a season characterised by lower temperatures and higher precipitation and humidity.

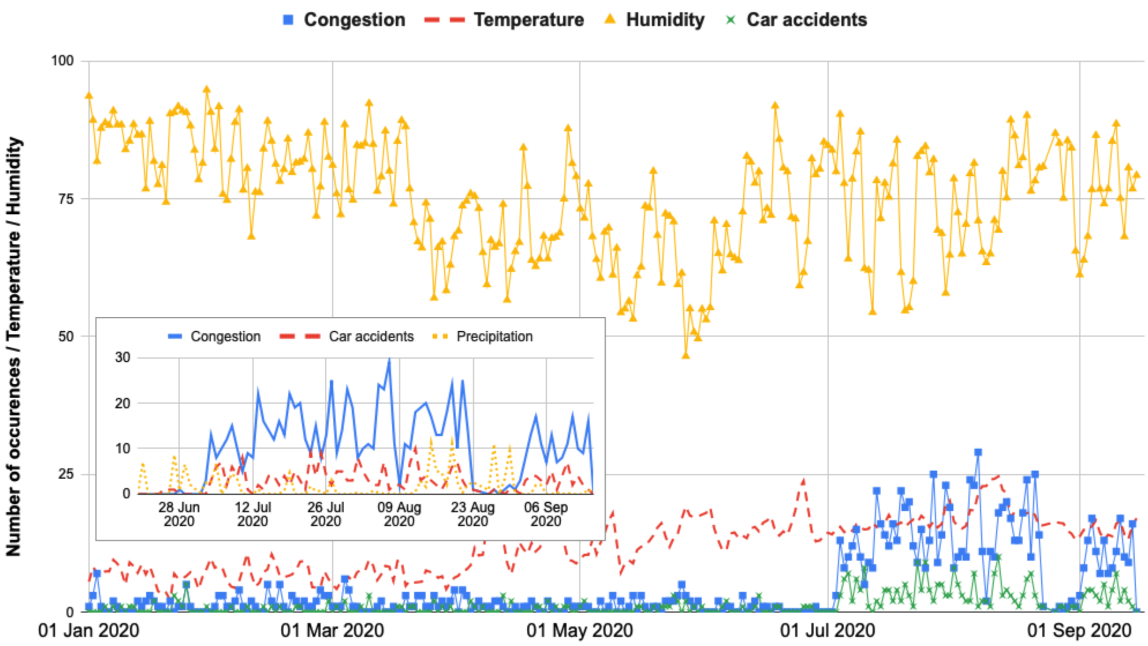
6.2.2 Scenario 1: Dependence of car accidents on congestion and environmental factors

In the first part of the current scenario (Scenario 1.1), we considered *car accidents* as a dependent variable and *congestion*, *temperature* and *humidity* as independent variables. In the second part (Scenario 1.2), we replaced *humidity* with *precipitation* as a dependent variable.

6.2.2.1 Preliminary results

From Figure 36, we can infer a general *positive correlation* between *congestion*, *humidity* and *car accidents* throughout the evaluated period, with peaks in activity beginning in early July and ending in early-mid September.

Figure 36: Preliminary results of social media data analysis for Scenario 1 - Smart City event occurrences



This trend could be explained by the summer period, generally accompanied by increased activity on the road due to the summer holiday season. In general, days with *peaks* in congestion and humidity also have *peaks* in car accidents. This occurrence repeats for days with *drops* in both metrics. We can infer a general *positive correlation* between *congestion*, *humidity* and *car accidents*.

6.2.2.2 Regression Statistics and ANOVA - Scenario 1.1

Table 7: Regression statistics for Scenario 1.1

Multiple R	0.62
$\mathbf{R^2}$	0.39
Adjusted $\mathbf{R^2}$	0.38
Standard Error	1.51

Table 8: ANOVA Table for Scenario 1.1

	SS	*MS*	*F*	*Sig. F*
Regression	371.79	123.93	54.05	4.92E-27
Residual	580.03	2.29		
Total	951.82			

In Table 8, $R^2 = 0.39$ suggests a low but positive intensity connection between variables. *Adjusted* $R^2 = 0.38$ shows that 38% of the variation of the dependent variable (*car accidents*) can be explained by the simultaneous variation of the independent variables (*congestion*, *temperature* and *humidity*). The small value of $f - test = 4.92231E - 27$ leads us to reject H_0 for the risk of 1% and 5% and indicates that there is a linear statistical dependence of medium intensity between variables.

6.2.2.3 Regression coefficients and variability - Scenario 1.1

Table 9: Regression coefficients for Scenario 1.1

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-2.192666185	0.8666880374	-2.52993706	0.01201601416
Congestion	0.1466703356	0.01793238305	8.179076658	1.39E-14
Temperature	0.09212658428	0.02417168049	3.811343789	0.000173632911
Humidity	0.02302590679	0.009790444796	2.351875453	0.019446223

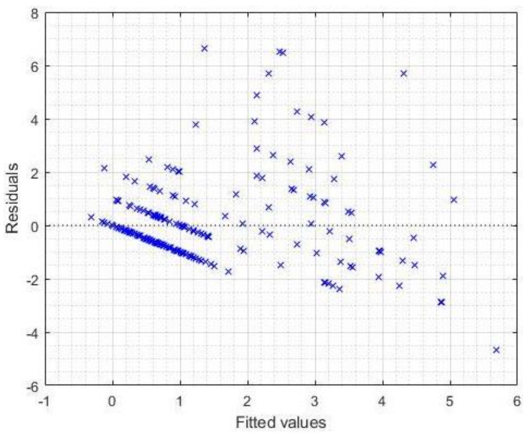
Table 9 dictates the equation of the multiple linear regression model: $Y = -2.192 + 0.146 * x_1 + 0.092 * x_2 + 0.23 * x_3$, where x_1 = *congestion*, x_2 = *temperature*, and x_3 = *humidity*.

After computing the *t-tests*, we reject H_0 due to $\sigma < \alpha$. All parameters are statistically significant. For β_0 (*car accidents*), $\frac{b_0}{s_{\hat{\beta}0}} = -2.192/0.86 = -2.5$. For β_1 (*congestion*), $\frac{b_1}{s_{\hat{\beta}1}} = 0.146/0.017 = 8.6$. For β_2 (*temperature*), $\frac{b_2}{s_{\hat{\beta}2}} = 0.092/0.024 = 3.83$. For β_3 (*humidity*), $\frac{b_3}{s_{\hat{\beta}3}} = 0.023/0.0097 = 2.37$.

For β_0 (*car accidents*), the mean is -2.192 when all other variables are simultaneously equal to 0. For β_1, it increases by 0.1466 when *congestion* increases by *1 unit*. For β_2, it increases by 0.092 when *temperature* increases by 1 degree. For β_3, it increases by 0.023 when *humidity increases* by *1 unit*. The coefficient formula also confirms that the *congestion* variable has the strongest positive effect on the dependent variable (*car accidents*), while *humidity* has the least positive effect.

6.2.2.4 Residual analysis - Scenario 1.1

Figure 37: Plot of residuals vs fitted values - Scenario 1.1



The plot of the residual distribution over the fitted values in Figure 37 suggests a clear conical pattern emerging, which could be narrowed to missing variables in the model or dependencies present in the independent variables (*humidity* could be a function of *temperature*). However, a large proportion of the variation of the dependent variables can still be explained by the selected independent variables, suggesting a relatively significant model that deserves improvements.

6.2.2.5 Regression Statistics and ANOVA - Scenario 1.2

This Section commences the second part of the current scenario analysis, with the independent variable interchanges, as described at the start of Section 6.2.2.

Table 10: Regression statistics for Scenario 1.2

Multiple R	0.68
R^2	0.46
Adjusted R^2	0.45
Standard Error	1.41

Table 11: ANOVA Table for Scenario 1.2

	SS	*MS*	*F*	*Sig. F*
Regression	359.19	119.73	51.11	7.34E-26
Residual	592.63	2.34		
Total	951.82			

In Table 10, $R^2 = 0.46$ suggests a direct connection between the four variables, of *medium* intensity. *Adjusted* $R^2 = 0.45$ shows that 45% of the variation of *car accidents* variable can be explained by variations of *faulty lights* and *electricity charges*. The low value of f-test = 7.34212E−26 in Table 11 leads us to reject H_0 for the risk of 1% and 5% and declares that the model significantly explains the connection between the variables, exhibiting linear statistical dependence.

6.2.2.6 Regression coefficients and variability - Scenario 1.2

Table 12: Regression coefficients for Scenario 1.2

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-0.2623316889	0.2746086367	-0.9552929289	0.3403414577
Congestion	0.1546656002	0.01779547431	8.691288442	4.57E-16
Temperature	0.07181207595	0.0229204853	3.133095788	0.001933427905
Precipitation	0.004672604321	0.0257001739	0.1818121674	0.8558757174

Table 12 dictates the equation of the multiple linear regression model: $Y = -0.2623 + 0.1546 * x_1 + 0.0718 * x_2 + 0.0046 * x_3$, where x_1 = *congestion*, x_2 = *temperature*, and x_3 = *precipitation*.

For β_0 (*car accidents*), since $\frac{b_0}{s_{\hat{\beta 0}}} = 0.3277/0.1095 = 2.992$ and $\sigma > \alpha$, we do not reject H_0 (β_0 is not statistically significant). All other parameters are statistically significant, as we reject H_0 due to $\sigma < \alpha$ in the *t-tests*. For β_1 (*congestion*), $\frac{b_1}{s_{\hat{\beta 1}}} = 0.3277/0.1095 = 2.992$. For β_2 (*temperature*), $\frac{b_2}{s_{\hat{\beta 2}}} =$

0.3277/0.1095 = 2.992. For β_3 (*precipitation*), $\frac{b_3}{s_{\hat{\beta}3}} = 0.0046/0.025 = 0.18$. For β_0 (*car accidents*), the average is -0.26233 when the other variables are simultaneously equal to 0. For β_1, it increases by 0.154 when *congestion* increases by *1 unit*. For β_2, it increases by 0.07 when *temperature* increases by *1 unit*. For β_3, it increases by 0.0046 when *precipitation* increases by *1 unit*.

As *precipitation* does not have a significant statistical value for the model (based on the regression coefficients formula), we consider it unfeasible to reliably plot the residuals or perform a further analysis. In general, there is no significant difference compared to Scenario 1.1. The model containing the *humidity* variable appears stronger than the one containing the *precipitation* variable.

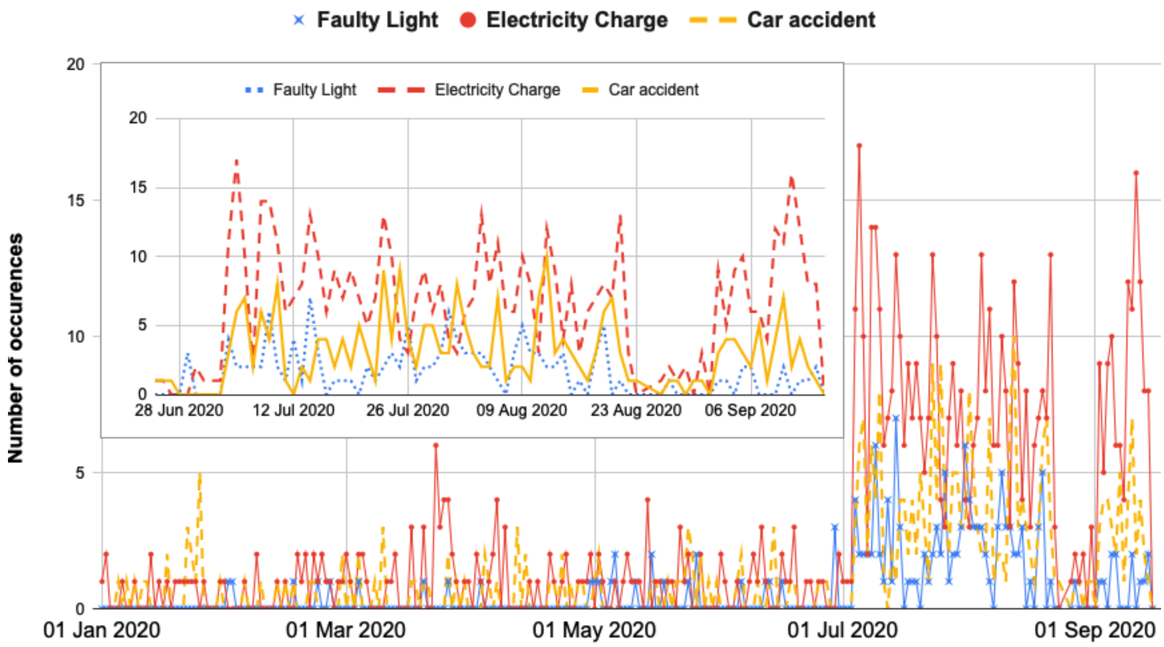
6.2.3 Scenario 2: Dependence of car accidents on faulty infrastructure

For this scenario, the researcher has used *car accidents* as a dependent variable, with *faulty lights* and *electricity charges* as independent variables.

6.2.3.1 Preliminary results

From Figure 38, we can infer a general *positive correlation* between *faulty lights*, *electricity charges* and *car accidents* throughout the evaluated period, with peaks in activity starting in early July and ending in early-mid September.

Figure 38: Preliminary results of social media data analysis for Scenario 2 - Smart City event occurrences



We could give the same explanation for this trend as in the previous scenario. In general, days characterised by *peaks* of *faulty lights* and *electricity charges* also feature *peaks* in *car accidents*. The same happens for days with *dips* in both metrics. In general, we can infer a general *positive correlation* between *faulty lights*, *electricity charges*, and *car accidents*.

6.2.3.2 Regression statistics and ANOVA

Table 13: Regression Statistics for Scenario 2

Multiple R	0.68
R^2	0.462
Adjusted R^2	0.458
Standard Error	1.41

Table 14: ANOVA Table for Scenario 2

	SS	*MS*	*F*	*Sig. F*
Regression	440.46	220.23	109.39	5.39E-35
Residual	511.36	2.01		
Total	951.82			

In Table 13, $R^2 = 0.462$ suggests a *direct* and *medium intensity* connection between the three variables. *Adjusted* $R^2 = 0.458$ shows that 45.8% of the variation of the *car accidents* variable can be explained by the variation of the *faulty lights* and the *electricity charges* variables. The estimated value of the *Multiple Correlation Ratio* $\sqrt{\frac{ESS}{TSS}} = 0.6802$ indicates a strong link between the model variables. From Table 14, based on the small value of the f test = 5.39E -35, we reject hypothesis H_0 for the risk of 1% and 5%. It can be stated that the model significantly explains the connection between the variables and that there is a linear statistical dependence between the variables.

6.2.3.3 Regression coefficients and variability

Table 15: Regression coefficients for Scenario 2

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	0.3275576729	0.109528596	2.990613271	0.003057774143
Faulty Lights	0.2690306152	0.09128012226	2.947307788	0.003503706361
Electricity Charges	0.2979115757	0.02992479622	9.955341835	6.31E-20

Table 15 dictates the equation of the multiple linear regression model: $Y = 0.32755 + 0.26903 * x_1 + 0.2979 * x_2$, where x_1 = *faulty lights*, and x_2 = *electricity charges*. All parameters are statistically significant since, after computing the *t-tests*, we reject H_0 due to $\sigma < \alpha$. For β_0 (*car accidents*), $\frac{b_0}{s_{\hat{\beta}0}}$ = 0.3277/0.1095 = 2.992. For β_1 (*faulty lights*), $\frac{b_1}{s_{\hat{\beta}1}}$ = 0.3277 / 0.1095 = 2.992. For β_2 (*electricity charges*), $\frac{b_2}{s_{\hat{\beta}2}}$ = 0.3277 / 0.12095 = 2.992. For β_0 (*car accidents*), the average is 0.3275 when the other variables are simultaneously equal to 0. For β_1, it increases by 0.269 when *faulty lights* increase by *1 unit*. For β_2, it increases by 0.297 when the *electricity charges* increase by *1 unit*.

6.2.3.4 Residual and outliers analysis

Figure 39: Case order plot of leverage - Scenario 2

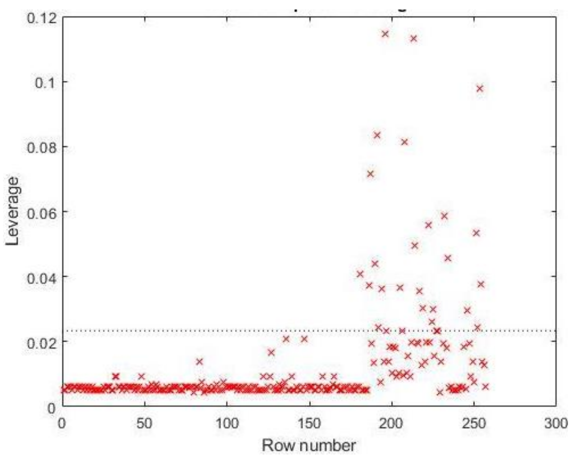
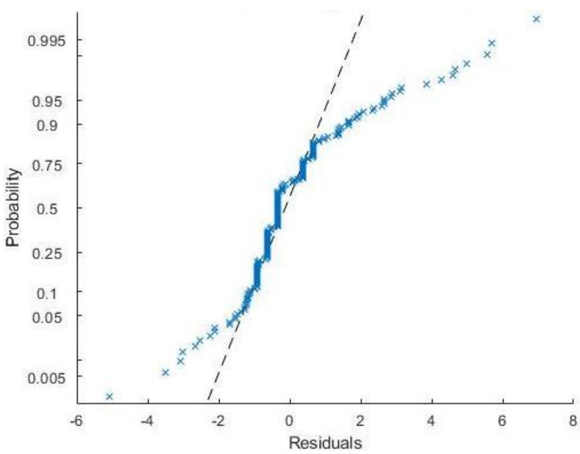


Figure 40: Probability plot of residuals - Scenario 2



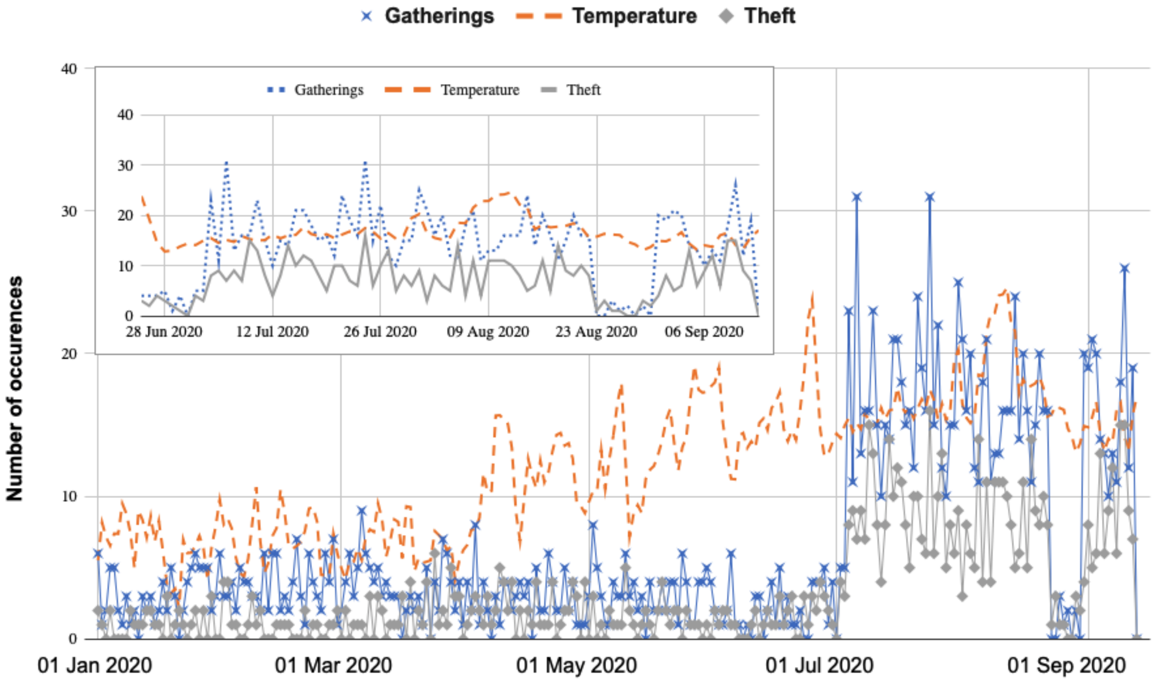
The residual analysis in Fig. 39 presents certain data sections with greater leverage than the rest, suggesting the presence of outliers, also confirmed by the standard distribution graph of residuals in Figure 40. A closer examination suggests that the residuals do not feature a normal distribution but rather a closer resemblance to a heavy tail. Although not a sufficient condition to invalidate the model, it presents the possibility of uncertainty within it. Overall, the model exhibits a reasonable degree of accuracy, despite the noticeable outliers present in the data. The model suggests that variations in the independent variable *faulty lights* have a greater impact than *electricity charges* on the dependent variable *car accidents*.

6.2.4 Scenario 3: Dependence of thefts on gatherings and temperature

In this scenario, we used *thefts* as a dependent variable, with *gatherings* and *temperature* as independent variables.

6.2.4.1 Preliminary results

Figure 41: Preliminary results of social media analysis for Scenario 3 - Smart City event occurrences



From Figure 41, we can infer a general *positive correlation* between *gatherings* and *thefts* throughout the evaluated period of 2020. We can observe a peak in activity starting from early July and spanning through early-mid September, which could be explained by the summer period, generally accompanied by more gatherings than the rest of the year, as well as higher temperatures. In general, days with *peaks* in *gatherings* also feature *peaks* in *thefts*. The same is true for days with *dips* in both metrics within a period that also coincides with higher *temperatures*. Overall, we can infer a general *positive correlation* between *gatherings*, *thefts*, and *temperature*.

6.2.4.2 Regression statistics and ANOVA

In Table 16, $R^2 = 0.64$ shows a direct link between the four variables of strong intensity. However, *Adjusted* R^2 0.64 means that 64% of the number of thefts can be explained by the variations of the *gatherings* and *temperature* variables.

Table 16: Regression statistics for Scenario 3

Multiple R	0.80
$\mathbf{R^2}$	0.64
Adjusted $\mathbf{R^2}$	0.64
Standard Error	2.21

Table 17: ANOVA Table for Scenario 3

	SS	*MS*	*F*	*Sig. F*
Regression	2290.21	1145.1	233.8	2.50E-5
Residual	1243.58	4.8		
Total	3533.79			

The value of the *Multiple Correlation Ratio* $\sqrt{\frac{ESS}{TSS}} = 0.805$ indicates a strong link between the variables. The small value of the f-test $= 2.4954\text{E}{-}58$ in Table 17 leads us to reject the hypothesis H_0 for the risk of 1% and 5%. We could state that the model significantly explains the connections between variables, exhibiting a linear statistical dependence.

6.2.4.3 Regression coefficients and variability

Table 18: Regression coefficients for Scenario 3

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-1.038999876	0.3673023976	-2.828731539	0.005046172269
Gatherings	0.3868833403	0.02363598937	16.3684005	1.32E-41
Temperature	0.1479463585	0.03215363987	4.601232057	0.000006635872262

Table 18 dictates the equation of the multiple linear regression model: $Y = -1.038 + 0.3868 * x_1 + 0.147 * x_2$, where $x_1 = gatherings$ and $x_2 = temperature$. All parameters are statistically significant since, after computing the *t-tests*, we reject H_0 due to $\sigma < \alpha$ in all $t - test$ values. For β_0 (*thefts*), $\frac{b_0}{s_{\hat{\beta}0}} = -\ 1.03\ /\ 0.36 = -\ 2.82$ For β_1 (*gatherings*), $\frac{b_1}{s_{\hat{\beta}1}} = 0.386\ /\ 0.023 = 16.38$. For β_2 (*temperature*), $\frac{b_2}{s_{\hat{\beta}2}} = 0.1479/0.032 = 4.60$. For β_0 (*thefts*), the average is equal to -1.03 when all other variables are simultaneously equal to 0. For β_1, it increases by 0.38 when *gatherings* increase by *1 unit*. For β_2, it increases by 0.14 when the *temperature* increases by *1 unit*.

6.2.4.4 Residual analysis

Figure 42: Residuals vs fitted values - Scenario 3

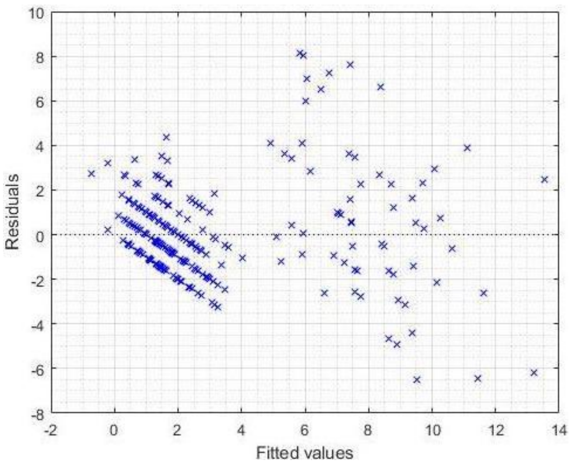
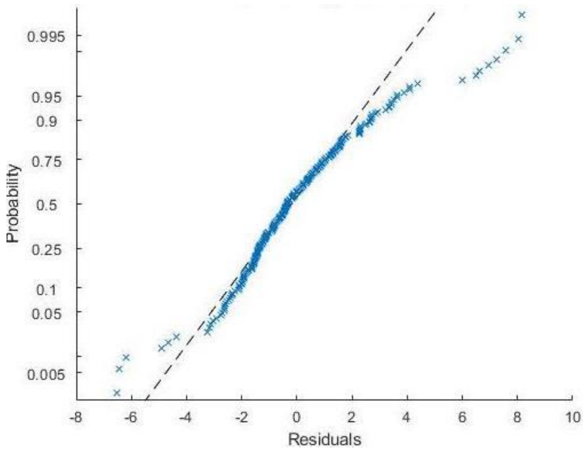


Figure 43: Normal probability of residuals - Scenario 3



The distribution of the residuals in Figure 42 exhibits a reasonable spread. The higher values suggest outliers present in the data, also observable in the standard probability plot (Figure 43), where deviation from the normal distribution is present in the lower and upper bands. Overall, the model is accurate despite noticeable outliers in the data. The coefficients formula confirms that both independent variables have a positive effect on the dependent variable *thefts*, while *gatherings* have a more pronounced influence than *temperature*.

6.3 Interlinks between Trends in Citizen Satisfaction and Concurrent Environmental Factors and Events

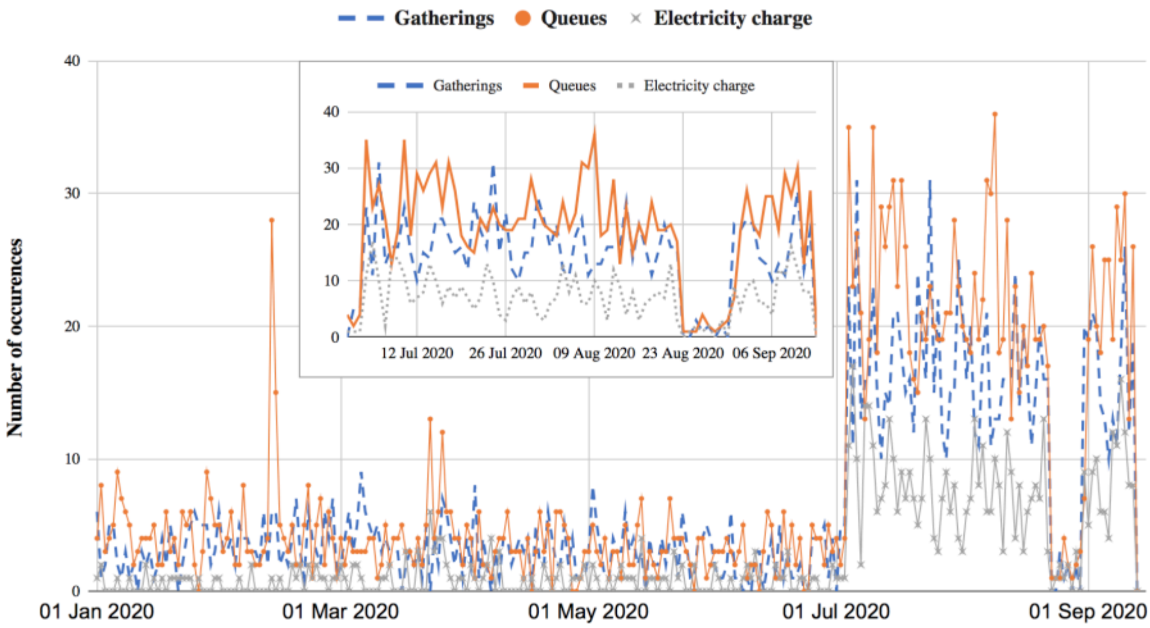
The final subsection highlights the intricate interconnections between citizen satisfaction trends and concurrent environmental factors and events in smart city contexts. A city is a vibrant, ever-changing ecosystem where environmental factors and significant events substantially shape residents' experiences. These factors can range from infrastructural developments, policy changes, and cultural events to more subtle changes in social and cultural dynamics. We explore how these variables interact with and influence trends in citizen satisfaction. This analysis sheds light on the dynamic interplay between citizens, their environment, and the city's semantic layer, providing a holistic view of how city-scale semantics influence user experiences.

6.3.1 Scenario 4: Dependence of citizen satisfaction with gatherings, queues, and electricity charges

In the first part of the current scenario analysis (Scenario 4.1), the researcher has used *positive sentiments* as a dependent variable and *gatherings*, *queues* and *electricity charges* as independent variables. In the second part (Scenario 4.2), the researcher has replaced the dependent variable with *negative sentiments*.

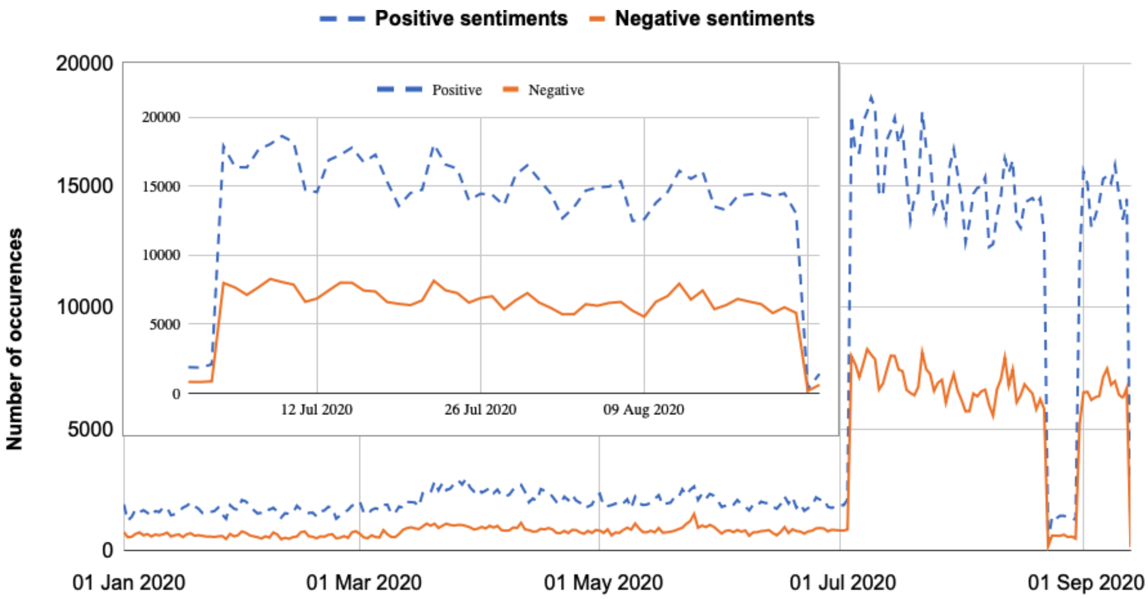
6.3.1.1 Preliminary results

Figure 44: Preliminary results of social media data analysis for Scenario 4 - Smart City event occurrences



When observing the preliminary results for Scenario 4 as a whole, Figure 44 indicates a general *positive correlation* between the emotional predilection of citizens and *gatherings*, with *peaks* of *positive sentiments* occurring in summer and within the proximity of periods characterised by low occurrences of *queues* and *electricity charges* respectively. Figure 45 shows a *peak* of *positive sentiments* in the second half of summer (July 25th to August 30th), followed by a *dip* in the first part of autumn (September 1st - 13th). This trend could be partially explained by the co-occurrence of COVID restrictions with the start of the autumn season, featuring higher levels of humidity and precipitation. We can also observe small *peaks* of *positive* emotional predilection in April and May 2020, periods with fewer *queues* and *electricity charges*. Marginal exceptions are days characterised by *peaks* in queues and *dips* in *gatherings*, which are also characterised by *dips* in *positive sentiments* (the days around 1^{st} of May and the days in late June).

Figure 45: Preliminary results of social media data analysis for Scenario 4 - Citizen satisfaction a Smart City



In general, we can infer a *negative correlation* between the independent variables (*queues* and *electricity charges*) and the independent variable (*positive sentiments*).

6.3.1.2 Regression statistics and ANOVA - Scenario 4.1

In Table 19, the first indicator of a valid and significant model is $R^2 = 0.92$, which shows a robust and direct link between variables. A more detailed analysis may raise the issue of noncompliance with the hypothesis regarding the modelling error, and the co-linearity phenomenon might occur.

Table 19: Regression statistics for Scenario 4.1

Multiple R	0.96
R^2	0.92
Adjusted R^2	0.92
Standard Error	1615.42

Table 20: ANOVA Table for Scenario 4.1

	SS	*MS*	*F*	*Sig. F*
Regression	7825425722	2608475241	999.57	6.31E-140
Residual	660225378.3	2609586.47		
Total	8485651100			

Adjusted $R^2 = 0.921$ means that the simultaneous variation of the independent variables can explain 92.1% of the variation of the positive variable. The estimated value of the *Multiple Correlation Ratio* $\sqrt{\frac{ESS}{TSS}} = 0.956$ indicates a strong link between the model variables.

In Table 20, the small f-test =6.306E−140 leads us to reject the hypothesis H_0 for a risk of 1% and a 5% risk. It can be stated that the model significantly explains the connection between the variables, and a *linear statistical dependence* exists between the variables.

6.3.1.3 Regression coefficients and variability - Scenario 4.1

Table 21: Regression coefficients for Scenario 4.1

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-69.3800431	142.6485889	-0.4863703428	0.6271256406
gatherings	339.1351667	29.05710087	11.67133529	1.82E-25
Queues	220.5609509	24.23102201	9.102420479	2.73E-17
Electricity Charges	465.1377969	48.73416334	9.54438868	1.23E-18

Table 21 dictates the equation of the multiple linear regression model: $Y = -69.380 + 339.135 * x_1 + 220.5 * x_2 + 465.13 * x_3$, where x_1 = *gatherings*, x_2 = *queues*, and x_3 = *electricity charges*. For the parameter *positive sentiments*, we do not reject H_0 and conclude that the parameter β_0 is not statistically significant, since $\frac{b_0}{s_{\hat{\beta}0}} = -69.38/142.6 = -0.48$ and $\sigma > \alpha$. All other parameters are statistically significant, since after computing the *t tests*, we reject H_0 due to $\sigma < \alpha$.

For β_1 (*gatherings*), $\frac{b_1}{s_{\hat{\beta}1}} = 339.13/29.05 = 11.67$. For β_2 (*queues*), $\frac{b_2}{s_{\hat{\beta}2}} = 220.56/24.23 = 9.10$. For β_3 (*electricity charges*), $\frac{b_3}{s_{\hat{\beta}3}} = 465.13/48.73 = 9.54$. For β_0 (*positive sentiments*), the average is -69.38 when all other variables are simultaneously equal to 0. For β_1, it increases by 399.135 when the number of *gatherings* increases by *1 unit*. For β_2, it increases by 220.5 when *queues* increase by *1 unit*. For β_3, it increases by 465.137 when *electricity charges* increase by *1 unit*. The coefficient formula also confirms that *gatherings*, *queues* and *electricity charges* have a positive effect on the dependent variable *positive sentiments*, with *electricity charges* having a higher positive effect.

6.3.1.4 Residual analysis - Scenario 4.1

Figure 46: Residuals vs Observation Order Scenario 4.1

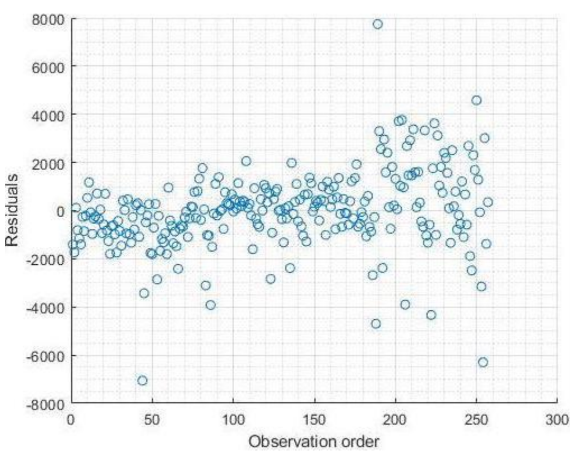
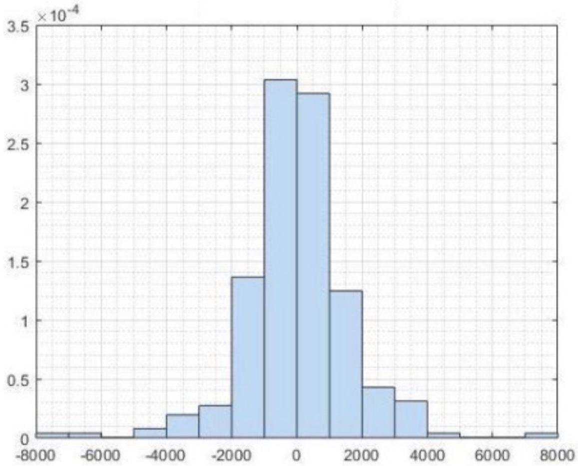


Figure 47: Histogram of residuals for Scenario 4.1



The plot in Figure 46 illustrates the residual distribution against the observation order. The primarily linear upper and lower bands suggest a significant relationship between the independent and dependent variables. Outliers can be observed in the residual histogram (Figure 47), which also shows that most residuals follow the normal distribution. Despite the noticeable presence observed around the range of 6000 – 8000, the model is accurate and representative.

6.3.1.5 Regression statistics and ANOVA - Scenario 4.2

This Section commences the second part of the scenario, where we considered *negative sentiments* as a dependent variable and *gatherings*, *queues* and *electricity charges* as independent variables.

Table 22: Regression statistics for Scenario 4.2

Multiple R	0.95
R^2	0.91
Adjusted R^2	0.91
Standard Error	774.07

Table 23: ANOVA Table for Scenario 4.2

	SS	*MS*	*F*	*Sig. F*
Regression	1621137835	540379278.2	901.84	9.76E-13
Residual	151596626.4	599196.15		
Total	1772734461			

The model with the *negative* dependent variable is almost as strong as the one with the *positive* dependent variables. From Table 22 we can observe a strong and positive correlation, where the value of *Adjusted R^2*, 91.3% of the variation of the *negative* variable can be explained by the variation of the independent variables in the model. The estimated value of the *Multiple Correlation Ratio* $\sqrt{\frac{ESS}{TSS}} = 0.956$ indicates a strong link between the model variables. The low value of f-test

=6.306E−140 in Table 23 leads us to reject H_0 for both 1% and 5% risk and state that the model significantly explains the connection between the variables, exhibiting a linear statistical dependence.

6.3.1.6 Regression coefficients and variability - Scenario 4.2

Table 24 dictates the equation of the multiple linear regression model: $Y = -128.47 + 148.63 * x_1 + 102.27 * x_2 + 217.68 * x_3$, where x_1 = *gatherings*, x_2 = *queues*, and x_3 = *electricity charges.*

Table 24: Regression coefficients - Scenario 4.2

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-128.471671	68.35434741	-1.879495245	0.06132622005
gatherings	148.6386076	13.92358089	10.67531469	3.31E-22
Queues	102.2776672	11.61102054	8.808671631	2.06E-16
Electricity Charges	217.6853553	23.35243521	9.321741109	5.90E-18

For β_0 (*negative sentiments*), since $\frac{b_0}{s_{\hat{\beta 0}}} = -128.47/68.35 = -1.87$ and $\sigma > \alpha$, we do not reject H_0 and conclude that the parameter β_0 is not statistically significant. All other parameters are statistically significant since, after computing the *t tests*, we reject H_0 due to $\sigma < \alpha$. For β_1 (*faulty lights*), $\frac{b_1}{s_{\hat{\beta 1}}} = 148.63/13.92 = 10.6$. For β_2 (*queues*), $\frac{b_2}{s_{\hat{\beta 2}}} = 102.2/11.6 = 8.80$. For β_3 (*electricity charges*), $\frac{b_3}{s_{\hat{\beta 3}}} = 217.68/23.35 = 9.31$.

For β_0 (*negative sentiments*), the average is -128.47 when all other variables are simultaneously equal to 0. For β_1, it increases by 148.63 when the number of gatherings increases by *1 unit.* For β_2, it increases by 102.27 when the number of queues increases by *1 unit.* For β_3, it increases by 217.68 when the *electricity charges* increase by *1 unit.* The coefficient formula also confirms that *gatherings*, *queues* and *electricity charges* have a positive effect on the dependent variable, with *electricity charges* having a higher positive effect.

6.3.1.7 Residual analysis - Scenario 4.2

Figure 48 illustrates the residual distribution against the observation order, with mostly linear upper and lower bands suggesting a significant relationship between the independent dependent variables. Despite the outliers present in the residual histogram (Figure 49), a normal distribution (-3000 - 3000) indicates an accurate and representative model.

Figure 48: Residuals vs observation order - Scenario 4.2

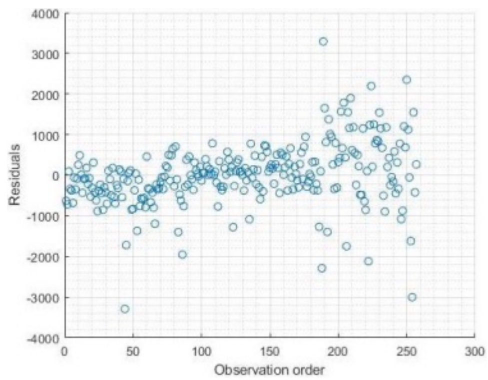
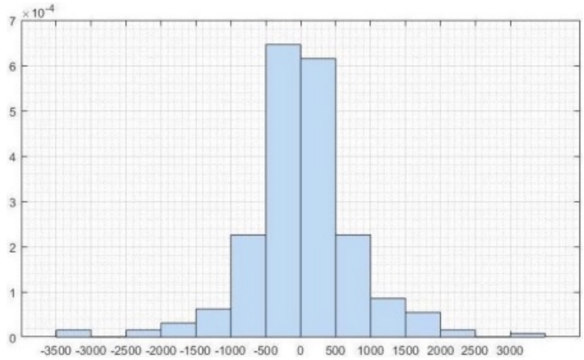


Figure 49: Histogram of residuals - Scenario 4.2



6.3.2 Scenario 5: Dependence of citizen satisfaction on environmental factors

In the first part of this scenario analysis, we considered *positive sentiments* as a dependent variable and *temperature*, *precipitation* and *humidity* as independent variables. In the second part, the dependent variable is replaced by *negative sentiments*.

6.3.2.1 Preliminary results

Figure 50: Preliminary results of social media analysis for Scenario 5 - Smart City event occurrences

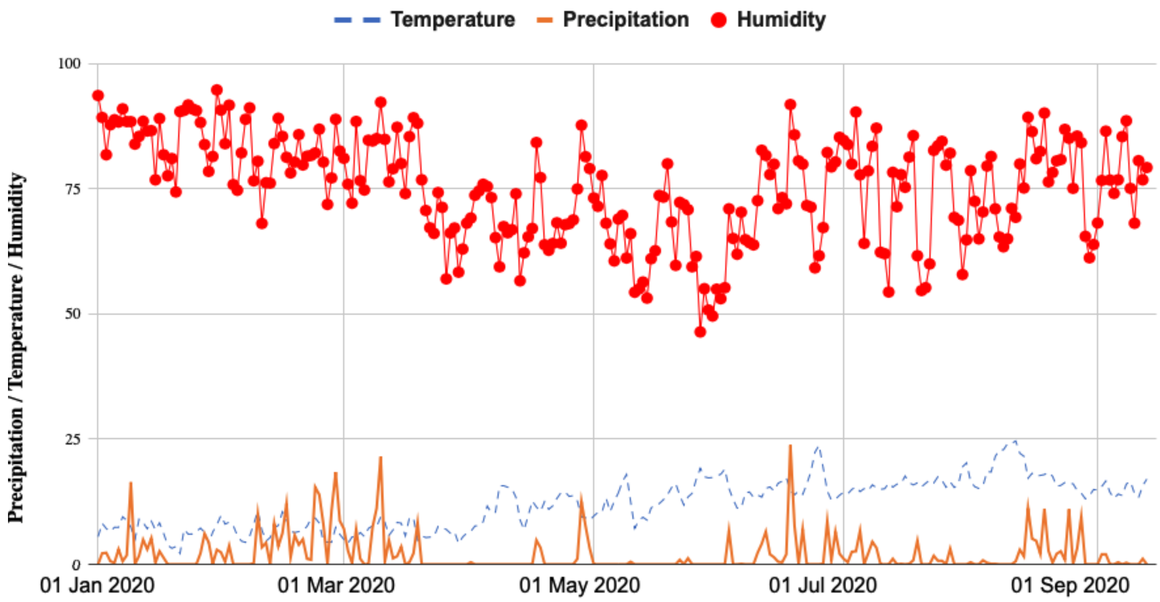
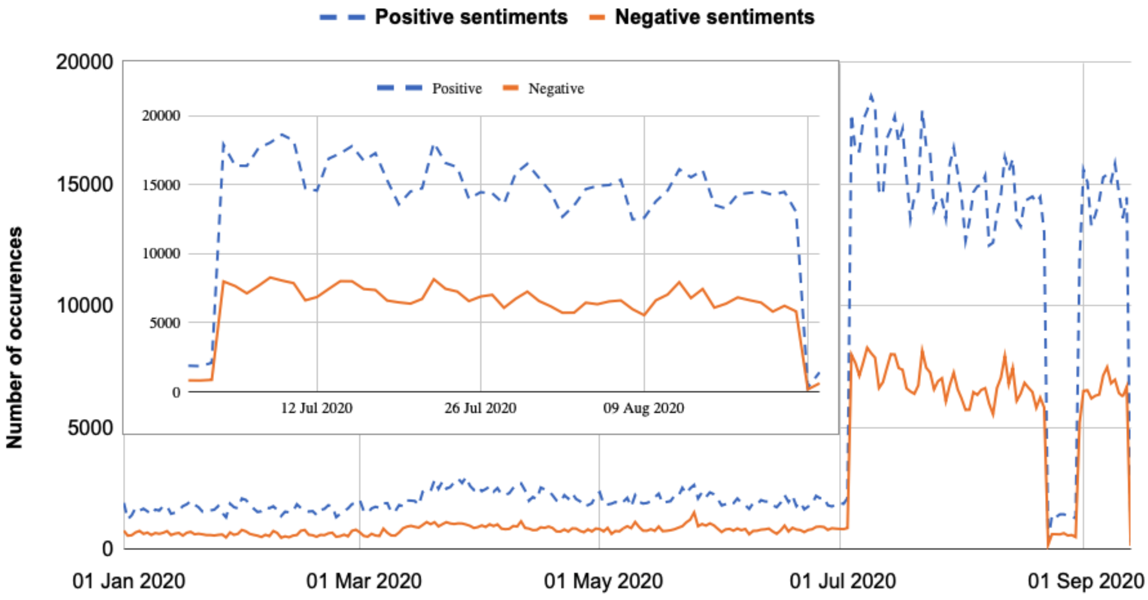


Figure 51: Preliminary results of social media analysis for Scenario 5 - Citizen satisfaction in a Smart City Context



We can observe *peaks* of positive sentiments in the second half of summer (July 25^{th} to August 30^{th}), coupled with a *dip* in the first part of autumn (September 1^{st}-13^{th}), which could be explained by lockdowns that occur together with the start of autumn, with high *precipitation* and *humidity*.

As in other scenarios, we can observe small *peaks* of positive emotional predilection in April and May 2020, months with *higher temperature* and *lower precipitation*. Marginal exceptions are days with *peaks* in precipitation and *dips* in temperature, which are also characterised by *dips* in positive sentiments (days around 1^{st} of May and days in late June). A general *positive correlation* emerges between citizen satisfaction and weather, with *peaks* of positive sentiments occurring in summer and close to periods featuring low *precipitation* and *humidity*.

6.3.2.2 Regression statistics and ANOVA - Scenario 5.1

Table 25: Regression statistics for Scenario 5.1

Multiple R	0.60
R^2	0.36
Adjusted R^2	0.35
Standard Error	4616.48

Table 26: ANOVA Table for Scenario 5.1

	SS	*MS*	*F*	*Sig. F*
Regression	3093729675	1031243225	48.38	9.45E-25
Residual	5391921425	21311942.39		
Total	8485651100			

From Table 25, we can observe potential improvements for the model, since $R^2 = 0.364$ suggests a direct connection, but of low intensity: only 35.7% of the variation of *positive sentiments* can be explained by the variation of the three independent variables. *Multiple Correlation Ratio* $\sqrt{\frac{ESS}{TSS}} = 0.603$ indicates an average link between the model variables. Observing Table 26, the low value of f-test $= 9.4516E - 25$ leads us to reject H_0, for both 1% and 5% risk. Therefore, the model significantly explains the connections between variables and exhibits linear statistical dependence.

6.3.2.3 Regression coefficients - Scenario 5.1

Table 27 dictates the equation of the multiple linear regression model: $Y = -12262.061 + 733.92 * x_1 - 201.87 * x_2 + 120.06 * x_3$, where x_1 = *temperature*, x_2 = *precipitation*, and x_3 = *humidity*. All parameters are statistically significant, since, after computing the *t tests*, we reject H_0 due to $\sigma < \alpha$.

Table 27: Regression coefficients and variability - Scenario 5.1

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-12262.06101	2713.834916	-4.518351847	0.000009571525336
Temperature	733.9261444	62.19320143	11.80074554	6.76E-26
Precipitation	-201.8790636	86.32112974	-2.338698117	0.02013027635
Humidity	120.0656664	32.63285245	3.679288123	0.0002857606397

For β_0 (*positive sentiments*), $\frac{b_0}{s_{\hat{\beta 0}}} = -12262/2713 = -4.518$. For β_1 (*temperature*), $\frac{b_1}{s_{\hat{\beta 1}}} = 733/62 = 11.8$. For β_2 (*precipitation*), $\frac{b_2}{s_{\hat{\beta 2}}} = -201/86 = -2.33$. For β_3 (*humidity*), $\frac{b_3}{s_{\hat{\beta 3}}} = 120/32 = 3.67$.

For β_0 (*positive sentiments*) the average is -12262.061 when all other variables are simultaneously equal to 0. For β_1, it increases by 733.92 when *temperature* increases by *one degree*. For β_2, it decreases by -201.87 when *precipitation* increases by *1 unit*. For β_3, it increases by 120.06 when *humidity* increases by *1 unit*. The coefficient formula also confirms that both *temperature* and *humidity* have a positive effect on the dependent variable, *temperature* having the highest. On the contrary, *precipitation* has a negative impact on *positive sentiments*.

6.3.2.4 Residual analysis - Scenario 5.1

The distribution of the residuals in Figure 52 presents a clear pattern in the data. The downward pointing line distribution plot suggests that a linear model is not the optimal representation of the data, or that clear dependencies are present in the model data, potentially explained by the physical interpretation of the variables: *humidity* could be a function of *temperature*, *precipitation*, or both.

Figure 52: Residual vs fitted values - Scenario 5.1

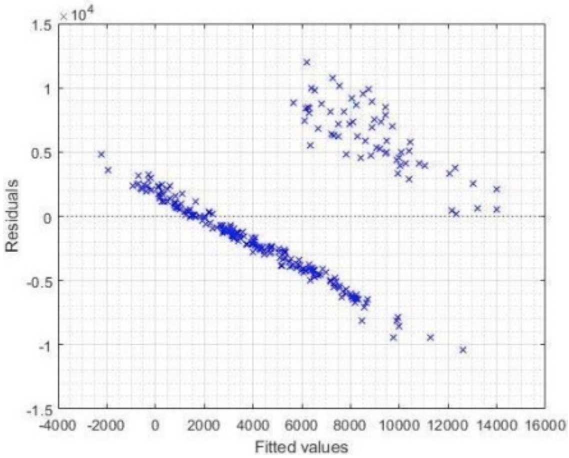
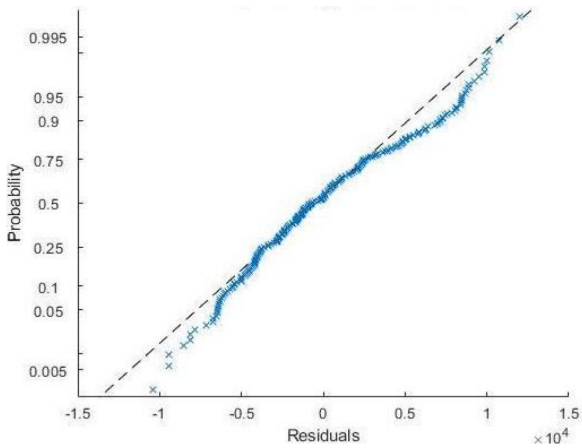


Figure 53: Normal probability of residuals - Scenario 5.1



The standard probability graph in Figure 53 suggests that although noticeable outliers are present in the data, the overall distribution is normal. However, as the distribution of the residuals over the fitted values indicates, the independent variables are not, in fact, entirely independent.

6.3.2.5 Regression statistics and ANOVA - Scenario 5.2

This section begins the second part of the scenario analysis, where we considered *negative sentiments* as a dependent variable and *temperature*, *precipitation*, and *humidity* as independent variables.

Table 28: Regression stat. - Scenario 5.2

Multiple R	0.61
R^2	0.37
Adjusted R^2	0.37
Standard Error	2086.85

Table 29: ANOVA Table for Scenario 5.2

	SS	*MS*	*F*	*Sig. F*
Regression	670931181.5	223643727.2	51.35	5.88E-26
Residual	1101803280	4354953.674		
Total	1772734461			

From Table 28, we can infer that the model could benefit from improvements, as $R^2 = 0.37$ suggests a direct connection, but of low intensity, and only 37.1% of the variation of the dependent variable can be explained by the variation of the three independent variables. The value of the *Multiple Correlation Ratio* $\sqrt{\frac{ESS}{TSS}} = 0.615$ indicates an average intensity link between the model variables. The small value of the f-test $=5.879E{-26}$ in Table 29 leads us to reject H_0 for a risk of 1% and 5%. It can be stated that the model significantly explains the connection between the variables, and a linear statistical dependence exists between the variables.

6.3.2.6 Regression coefficients and variability - Scenario 5.2

Table 30 dictates the equation of the multiple linear regression model: $Y = -5871.49 + 342.52 * x_1 - 89.12 * x_2 + 55.72 * x_3$, where x_1 = *temperature*, x_2 = *precipitation*, and x_3 = *humidity.*

Table 30: Regression coefficients for Scenario 5.2

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-5871.495358	1226.771502	-4.786136087	0.000002893788971
Temperature	342.5242927	28.11403402	12.18339184	3.55E-27
Precipitation	-89.12472999	39.02090778	-2.284025028	0.02319907252
Humidity	55.7241524	14.75146966	3.777532252	0.0001975163616

Table 30 also suggests that all parameters are statistically significant since the calculation of the *t tests* leads us to reject H_0 due to $\sigma < \alpha$.

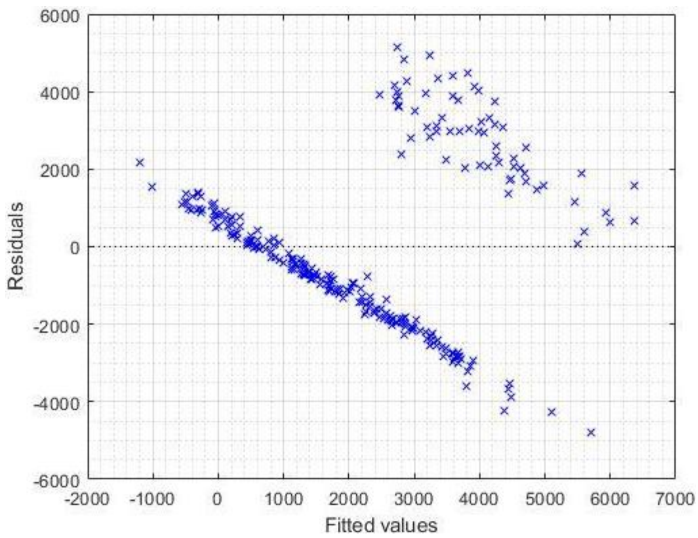
For β_0 (*negative sentiments*), $\frac{b_0}{s_{\hat{\beta 0}}} = -5871.4/1226.7 = -4.78$. For β_1 (*temperature*), $\frac{b_1}{s_{\hat{\beta 1}}} = 342.52/28.11 = 12.18$. For β_2 (*precipitation*), $\frac{b_2}{s_{\hat{\beta 2}}} = -89.12/39.02 = -2.2$. For β_3 (*humidity*), $\frac{b_3}{s_{\hat{\beta 3}}} = 55.72/14.75 = 3.77$. For β_0 (*negative sentiments*), the average is -5871.4 when all other variables are simultaneously equal to 0. For β_1, it increases by 342.52 when *temperature* increases by one degree. For β_2, it decreases by -89.12 when precipitation increases by one unit. For β_3, it increases by 55.72 when humidity increases by *1 unit.*

The coefficient formula also confirms that *temperature* and *humidity* positively impact the dependent variable (*negative sentiments*), with *temperature* having a higher positive impact. We could also note that *precipitation* has a negative impact on the dependent variable (*negative sentiments*).

6.3.2.7 Residual analysis - Scenario 5.2

Figure 54 represents the distribution plot of the residuals in the fitted values. Although appearing similar to Scenario 5.1, it also suggests that the independent variables are not, in fact, entirely independent. However, both models are promising. Independent variables largely explain the variation of the dependent variable. We could notice a higher accuracy in the model featuring *positive sentiments* as a dependent variable than its counterpart. Part of future work plans is to treat the colinearity problem while testing the hypothesis of errors.

Figure 54: Plot of residuals vs fitted values - Scenario 5.2



In this chapter, the researcher used natural language processing (NLP) techniques to analyse social media data, providing crucial information for decision making. The researcher applied these techniques to Twitter messages, enabling automatic event detection in real time. To effectively classify events, the researcher created a risk taxonomy, facilitating the identification of related events from data streams and attributing citizen satisfaction values to them.

We optimised the language processing models by striking a balance between accuracy and processing speed and compared various models previously used in research for event classification. This resulted in an impressive accuracy of 88.5%. For a real-world context, the researcher examined several smart city and community initiatives.

Using Multiple Regression Analysis (MRA), we found a significant correlation between different types of events and citizen satisfaction, considering environmental factors as well. The MRA models revealed a variance of 60% - 90% between the dependent and independent variables and included at least one variable with statistical significance (below the 0.05 threshold).

Our findings highlight the crucial role of citizens as active data providers in the development of smart cities. Therefore, this chapter improves understanding of citizen-city interactions, demonstrating the importance of semantic analysis in user experience within smart city contexts. These insights help improve web applications in urban settings, setting the stage for future smart city advancements.

6.4 Evolution of the bespoke UX model

At this stage of research, the custom UX model experienced a significant transformation. It was adapted to suit the unique challenges and opportunities of engaging a diverse, city-scale on-line social community. This stage diverged from the controlled environments of previous case studies, introducing complex factors such as scale, diversity, and indirect user engagement through social media content. The participatory design process changed, moving from direct user interviews to analysis of user-generated content on social networks, reflecting an adjustment in the engagement strategy appropriate for the wide-ranging urban environment.

Central to the model's evolution at this level was harnessing collective intelligence through social media. Citizens effectively became active sensors, providing semantic data through their interactions and postings, yielding valuable insights into urban life's fabric. Advanced Natural Language Processing (NLP) deployment allowed the model to extract meaning from this flood of data. It interpreted tangible aspects, such as traffic patterns and weather conditions, and more nuanced intangible aspects, such as social interactions and cultural subtleties.

Due to the vastness of the cityscape and the impracticality of conducting direct interviews, the UX model was adapted learning to handle location-specific real-time data sourced from social media. This adaptation was crucial, transforming the model into a more dynamic instrument that provided an ever-current reflection of the city's lifeblood. The transition to reliance on indirect user feedback collected from social networks helped develop urban environments that were not only responsive but also inclusive and flexible.

Navigating through the intricacies of scale and diversity, the model incorporated crowd-sensing applications to understand the context and facilitate connections between various roles of the system, such as citizens, developers, and stakeholders. Privacy issues were of particular concern due to the indirect nature of data collection and the sensitivity of the information. The model was adapted by prioritising user privacy and committing to ethical data handling practices that respect individual privacy rights.

Concerning its decision-making capabilities, the model developed an innovative conceptual framework using social media content as a surrogate for citizen engagement in smart cities. The framework suggested that by interpreting and responding to data produced by citizens, cities themselves could evolve to become more dynamic and responsive. This phase of the evolution also saw the development of a new taxonomy designed to classify urban events based on risk levels detected from social media data, enabling swift and informed decision-making processes without the need for direct user contact.

An essential aspect of the model was its focus on sentiment analysis of social media content, which became crucial in measuring satisfaction levels among citizens with ongoing events and environmental factors. This indirect method of interpreting the city's "pulse" provided an additional layer of understanding, positioning the model as a tool capable of sensing both the data-driven and emotional contours of urban living.

The evolution of the bespoke UX model throughout Stage III was characterised by its ingenious integration of data from an on-line social community on the scale of a city. This was achieved through advanced NLP, crowd sensing techniques, and creating a customised taxonomy that managed event classification. The model was developed to meet the challenges of scale and diversity, evolving into a dynamic and comprehensive tool capable of providing instant insights and supporting informed decision-making processes. The indirect participatory design procedure, appropriate for a city-wide context and driven by social media content, highlighted the model's versatility and promise to overhaul user experiences in smart city contexts. This stage underscored the criticality of ethical data practices, the maintenance of privacy, and the ongoing need to adapt to an ever-changing and diverse urban fabric, laying down benchmarks for future explorations and advancements at the crossroads of semantics, user experience design, building information modelling (BIM), and intelligent urban development.

7 Discussion and Evaluation

This section holistically examines the research carried out at the three stages, reflecting on the key findings, contributions, limitations, and potential future trajectories of the work. The research aims to demonstrate the value of applied semantic techniques in elevating user experiences within BIM web applications and the broader built environment context.

The three research questions framed in the Introduction have guided the exploration of this topic across progressive levels of scope and complexity. Stage I focused on improving learning in BIM using semantics, and Stage II used semantics to improve user interaction with an energy management platform. At the same time, Stage III encompassed scaling semantic frameworks to integrate citizen perspectives in urban digital twins.

Through literature reviews, prototype development, empirical studies, and expert feedback, each stage examines the convergence of semantics, user experience principles, and emerging technologies. The research methodology, underpinned by design science, pragmatism, and user-centredness, has facilitated a multidisciplinary approach to probe this convergence within training systems, energy analytics dashboards, and city-wide knowledge graphs.

This chapter of the *Discussion* provides a reflective and evaluative perspective on the results of the research. The key contributions and limitations of the original objectives are analysed, revealing critical insights.

Quantitative and qualitative methods will be used throughout the evaluation process. Surveys, interviews, usability tests and data analysis will be conducted to gather valuable information and assess the strengths and limitations of each approach. The results obtained from these evaluations will provide a comprehensive understanding of the effectiveness and impact of the proposed methodologies.

By evaluating the three approaches systematically and rigorously, this chapter contributes to the overall research objective of enhancing the user experience of BIM web applications through semantics. The evaluation results will validate the effectiveness of the proposed approaches and provide valuable information for further refinement and future research in semantics and user experience in the built environment.

7.1 Evaluating the enhancement of learning and training through semantics

The enhancement of learning and training through semantics, particularly within the domain of BIM, has been a focal point of this research. The bespoke UX model, intricately designed and meticulously refined, has been instrumental in this endeavour. The innovative approach of the model, using semantic technologies, has transformed learning experiences and training methodologies. This subsection evaluates the effectiveness of the model in improving learning and training, grounded in the insights derived from the BIMEET use case, the comprehensive questionnaire, and the thesis's overarching theoretical and practical framework.

The bespoke UX model, as detailed in the "Developing a new User Experience Model" subsection of the "Methodology" section, was centred around the principle of semantic enhancement. By integrating advanced semantic technologies and methodologies, the model sought to provide a more intuitive, efficient, and personalised learning experience for users engaged in BIM-related activities. The user-centric design of the model, underscored by the iterative feedback loop from various stakeholders, including learners, educators, and industry professionals, ensured that the model was not only technically sound but also aligned with its users' real-world needs and preferences.

The questionnaire, a vital component of the participatory design process, was crucial in shaping the UX model. The diverse array of questions aimed at understanding the field of expertise, the historical experience with BIM, and the specific skills required for energy efficiency in BIM provided invaluable insight into user learning and training needs. The responses highlighted the need for a UX model that was not only capable of delivering context-aware learning content but also adaptable to the varying levels of expertise and user learning styles.

Moreover, the BIMEET use case provided a practical context for applying and evaluating the UX model. Within this case study, the effectiveness of the model was tested in improving learning and training through semantics. The case study highlighted how the model facilitated the seamless integration of semantic web technologies into learning and training platforms, enabling users to access up-to-date, contextually relevant and personalised learning content. The model's ability to interpret and respond to user feedback in real time further enriched the learning experience, making it more engaging and effective.

It is essential to acknowledge the limitations encountered in this evaluation process. Although innovative, reliance on social media content and indirect user engagement also presented challenges to accurately capture and interpret the nuanced learning and training needs of users. Additionally, the relatively small sample sizes of users and experts involved in the feedback loop necessitated a cautious interpretation of the findings, highlighting the need for more extensive empirical studies to validate the model's effectiveness across broader populations.

This stage of the research provides validation of encouraging results in the context of the EU H2020 BIMEET project, in collaboration with the energy value chain of construction, as evidenced in the web portal (*energy-bim.com*).

The researcher has analysed a total of 40 million tweets sourced from a selection of Twitter accounts belonging to the friends and followers of the most prominent users of the *energy-bim.com* platform, as illustrated in Table 3. This corpus was later filtered down to "60,000 tweets" of relevance using regular expressions. To confirm the roles and competencies required by this training process, a scientific literature repository comprising "80 BIM publications" was analysed, mainly authored by the same organisations and connected authors.

The restricted size of the dataset may introduce a potential bias in the results obtained, as it was generated from a randomly selected sample of organisations within the Building Information Modelling (BIM) industry. Specifically, this sample comprises only those organisations that use the *energy-bim.com* portal, as well as their associated contacts and followers. This analysis could be further scaled for the entire BIM domain and for a larger representative sample of the construction industry by considering partners and organisations from other BIM collaboration projects, multiple countries and areas of expertise. This will have the potential to provide more information on training needs in the BIM industry.

Research could also benefit from the potential use of other text mining techniques in addition to TF-IDF and cluster analysis. As recently evidenced by [333], techniques such as NLP (Natural Language Processing) and classification are theoretically useful for deriving useful knowledge for educational stakeholders.

As part of the EU H2020 BIMEET project, the researcher has captured a database of training programmes provided by BIM accredited institutions that is actively updated, enabling institutions to continuously update their offerings (either scheduled or on-demand) with a detailed level of granularity. The training programmes are split into learning outcomes, prerequisites, and the levels of expertise expected after completion. Research findings can be associated with the learning outcomes described by organisations and recommendations for improvement could also be derived. These results could be further distributed to other stakeholders to encourage them to implement practical policies that consider the newly found correlations between roles and skills.

We acknowledge that the research does not incorporate emerging concepts in the industry, such as digital twins and scan-to-BIM. This is due to (a) the scope of the IFCs (Industry Foundation Classes) and (b) the sample used in the context of the investigation. In addition, this limitation can be overcome by (a) leveraging semantic associations using the two data sources (social media and scientific publications) and (b) scaling up the sample to include other third-party social media sources drawn from other disciplines and countries.

The authors have recently initiated follow-up research in the context of the H2020 INSTRUCT project to address the training needs of blue- and white-collar staff throughout the life cycle of a construction project. This will give us the opportunity to stress test the proposed text mining approach while benchmarking other text mining algorithms.

This study demonstrates that BIM is a complex undertaking that involves multiple disciplines and domains, from finance and management to engineering and technology, with significant implications for research and education. Based on such findings, we identify key industry trends that emphasise the competencies, skills, and roles required to implement more sustainable BIM practices for energy efficiency. Research methods are generic and can be used in wider research contexts to address the ongoing phenomenon of climate change.

The analysis has evidenced the importance of socio-organisational aspects in construction skills and the need to stimulate and sustain progress in the current and ongoing BIM implementation landscape. These findings corroborate the existing literature, in which these BIM positions and associated skills are generally overlooked. For example, most studies raised the awareness of an imminent need for vocational education to up-skill the construction workforce to address consistently and holistically the energy efficiency agenda. However, the construction industry is renowned for its resistance to change, which may hinder the current digitalisation agenda. As relevant correlation factors between BIM roles and skills were determined, it is worth highlighting that these correlations will change to reflect the continuous evolution of BIM roles and skills, demonstrating the incremental adoption and improvement in maturity of BIM in industry.

Furthermore, research has shown the role of BIM as an enabling technology for energy modelling. The use of BIM, i.e. the IFCs, falls short in describing complex energy systems. Here, the authors have augmented BIM with additional semantic resources, that is, ontologies, in domains such as water, energy and larger infrastructures, to represent complex artefacts. In fact, training for energy efficiency can be extended to include tools and software that make use of BIM, augmented with additional concepts drawn from these ontologies to perform various forms of engineering analysis, including energy modelling. In addition to the acknowledged role of social media, a change programme is necessary to fully embrace the digital transformation of the construction industry. Therefore, a holistic methodology was required, including additional sources of information, to assess BIM with associated competencies and training programmes.

Future research involves using a project-based approach to demonstrate the correlation between roles and skills by implementing the proposed methodology on real-world projects, thus providing tailored recommendations on the skills required for each identified project role. As such, the authors will organise further consultations with stakeholders involved in real-world projects and validate the assigned roles and skills, as well as the ensuing methodology, further informing the evolution of BIM training programmes delivered in the construction industry.

7.2 Evaluating the delivery of context-informed Digital Twins through semantics

The following sub-chapter presents a detailed evaluation of an energy performance solution implemented within a railway station, examined via the participatory action methodology within a workshop context. Using this approach, the aim was to involve various stakeholders in the evaluation and improvement of energy performance strategies, thereby facilitating a comprehensive understanding of the effectiveness of the intervention and potential areas for improvement.

Participatory action research (PAR) methodology, the cornerstone of this research approach, has been celebrated as an effective method of investigating complex systems in the academic literature [219]. Unlike traditional research methodologies, PAR encourages the participation of stakeholders at all stages of the research process, encouraging the collective exploration of issues and the collaborative creation and implementation of solutions [283].

In the context of evaluating energy performance solutions, the PAR approach is particularly pertinent. Railway stations, due to their diverse array of functions, stakeholders, and inherent complexities, represent challenging environments for energy performance interventions. The PAR methodology allows for integrating diverse perspectives - from station managers and operators to architects and engineers, to daily commuters and occasional users. These points of view can provide crucial information on real world challenges and opportunities, thus ensuring that the proposed solutions are not only technically viable, but also contextually appropriate and acceptable [283].

In essence, by applying a workshop format, participants are invited to reflect and critique the energy performance solution. This feedback is valuable in identifying unforeseen problems, potential improvements, and innovative ideas to further enhance energy performance [56]. This dynamic and collaborative approach supports more inclusive and sustainable outcomes, increasing the likelihood of successful implementation and adoption of the energy performance solution [228].

The following sub-chapter provides an in-depth investigation of the energy performance solution applied at a railway station, as evaluated through a workshop utilising the PAR methodology. This interactive approach seeks to foster a comprehensive understanding of the solution, ensuring that it meets the practical needs and expectations of those directly affected by its implementation.

7.2.1 Background of the evaluation

An important tech company, in collaboration with Cardiff University and the stakeholders of an important railway station in the UK, has initiated a Digital Twin of the station that is being used for the following purposes:

1. identifying opportunities for energy and carbon reduction
2. developing "What-If" scenarios to explore potential improvements
3. optimising the station's performance via energy efficiency measures and improved controls delivered via the Digital Twin.

For purpose (a), the digital twin of the railway station was used to identify opportunities for energy and carbon reduction by analysing data on the current energy use of the station, identifying areas where energy savings could be made, and creating simulations that model different scenarios of potential improvements. This step in the approach would allow railway operators to make more informed decisions about how to best reduce their energy consumption while maintaining a high level of service.

For purpose (b), the digital twin of the railway station was used to develop “What-If” scenarios to explore potential improvements by performing simulations that test different options to improve the efficiency or performance of the station. For example, this could include evaluating changes in schedule or operational strategies or introducing new technologies such as intelligent lighting systems. This operation would allow railway operators to identify which strategies are the most likely to succeed before investing in them.

For purpose (c), the digital twin of the railway station can be used to optimise performance through energy efficiency measures and improved controls delivered via the Digital Twin, facilitate real-time monitoring and control of energy use at different points in the network, and use predictive analytics to anticipate future demand and optimise energy usage accordingly. The digital twin would also enable railway operators to implement new technologies or initiatives quickly and effectively that could help improve performance efficiently and cost effectively.

7.2.2 Phase 1 - Establishing the energy profiling of the railway Station

Performing an Energy Audit of the Station

The first step was to conduct an energy audit of the station to understand and map:

1. The existing energy systems and how they are operated.
2. Energy-consuming equipment throughout the station and its operating schedule.
3. Existing detection and control infrastructure and their control systems.
4. User requirements (station users) and their associated behaviours.
5.

Consequently, the initial step consisted of the following:

- determined the location of the sensing nodes and the data they consume.
- validated models/drawings against as-built and station usage.
- determined whether occupancy data could be obtained from train schedules and any existing sensors.
- consulted facility and energy managers.
- ensured the dependability of local applications even in the absence of LAN/WAN connections.
- evaluated user behaviours using data from footfall, train schedules, sensors, and other sources, in addition to energy profiling, based on historical data.
- held meetings with the asset/energy manager to establish (up to three) scenarios and gain a deeper understanding of the asset.

A calibrated energy simulation model of a station was developed from an existing 3D CAD model of the building, along with data on occupancy, climate, and other factors. The model included details about the building envelope, mechanical systems, lighting fixtures, and other components. It should also include external influences, such as solar radiation, wind speed, and air temperature.

Once completed, the model was run through a simulation tool to generate results. These results will provide insight into how the station uses energy and how it could be improved for better efficiency. The model could then inform decisions about design changes that can lead to significant energy savings.

Based on the energy data and the BIM, an energy simulation model of the station was developed with associated recommendations regarding energy improvements based on the initial modelling. The final output consisted of the following:

1. An up-to-date BIM semantic model of the railway station includes all fabric and mechanical systems and equipment information.
2. A simulation model based on the information and assumptions gathered from the audit exercise.

7.2.3 Assumptions

The work was based on a number of assumptions:

1. Various inputs were required in addition to the information received prior to the commissioning of the evaluation.
2. The modelling/simulations were based on all available information post-audit.
3. The number of elements available from the list below was directly proportional to the accuracy of the model.

The elements required were the following:

1. BIM model for the railway station/area under the study of the station, together with CAD drawings, with complete elevations, sections and plan drawings for the station.
2. Construction materials / material properties of the existing structure.
3. Internal lighting details (number of luminaries and lamp efficiencies lm/W / installed lighting loads in W/m2) and lighting controls.
4. Lighting day lighting / dimming controls installed. Details of any metering systems installed.
5. Details of the station equipment (heat emissions, electrical power requirements, any working/standby arrangement). Electrical system schematic drawings.

6. Station occupancy statistics: number of people using the station, full-time occupants in an 'occupied area', and transient occupants per hour.
7. Station operational times throughout the year.
8. Details of installed HVAC systems, HVAC system drawings, HVAC equipment, and system capacities.
9. Details of installed hot water systems, drawings, equipment, and system capacities.
10. Class of ducts used in HVAC systems.

7.2.4 Objectives for Delivering an Energy Consumption Forecasting Capability

The first objective was to find the gaps in the current station operation. We wanted to identify any improvement areas and work to ensure that all passengers could have a safe and enjoyable trip. To achieve this, the researcher, along with the team, looked at the existing energy management practices and other aspects of the station's operations. We also conducted surveys with train station managers and collected feedback as part of the research process. By doing this, we could identify where we have actual data from the deal to use as a starting point where we would need to make assumptions concerning energy consumption.

Our second objective was to understand the delta difference between the ideal assumptions and the actual data related to energy consumption.

The delta difference between the ideal assumptions and the actual data related to energy consumption can be illustrated by looking at the discrepancies between the future and actual energy consumption projections in a given area. For example, if a model estimates that a particular region will consume 10 million kWh of electricity over a year, it will consume only 8 million kWh. The delta difference was 2 million kWh. This gap between expectations and reality can be attributed to changes in population, economic activity, or energy efficiency initiatives. By examining these discrepancies, we can better understand how assumptions about energy consumption may be flawed and adjust strategies accordingly.

Our third objective was to simulate a "what-if" scenario to find the best operating station settings.

We used the CUSP platform to host and manage the Digital Twin of the railway station and its associated engineering models, including an energy simulation model. The CUSP platform provided a secure and reliable environment to run the digital twin and its associated engineering models. It also enabled us to share the data from the Digital Twin with stakeholders in real time, allowing us to create an up-to-date overview of the station's energy consumption. CUSP also enabled us to integrate existing datasets into the digital twin, such as historical weather data, which helped inform the energy simulation model. Finally, CUSP allowed us to easily store and access the data collected from the Digital Twin.

The energy simulation model was calibrated on the basis of historical data collected over five years and live energy consumption data. The calibration involved analysing the historical energy consumption data to identify energy usage patterns and adjust the model parameters accordingly. This allowed the model to be calibrated to be more accurate in predicting future energy consumption. Live energy consumption data was also used to refine the model, allowing it to reflect current conditions with greater accuracy. Once the model had been calibrated, it could be used to make more accurate predictions about future energy use and develop strategies to reduce overall energy consumption.

We used the CUSP simulation data science service capabilities to generate random scenarios and tested the various settings for each scenario. We then used the CUSP optimisation data science service to identify the best settings for each scenario. This enabled us to determine which settings would perform best under different operating conditions. Finally, we used this information to develop a set of recommended operating parameters that could be used to optimise the station's performance in any given situation.

The model results should be analysed to ensure the accuracy and reliability of the predictions. Appropriate metrics (e.g., root mean squared error) should be used to evaluate the model's performance against observed values to identify areas where changes need to be made to improve the accuracy or reliability of predictions.

The model was tested with real-world data before being deployed into production environments to ensure the accuracy and reliability of the predictions in real-world scenarios. Additionally, measures must be taken for ongoing monitoring and maintenance of the system to ensure its continued correctness over time as new external factors come into play or existing ones change.

CUSP was used to define the baselines for energy consumption and carbon emissions and uncover sensitive variables that drive energy consumption and carbon emissions.

The future performance predictions were then developed using Deep Learning to predict the performance of the stations when interventions identified in the scenarios are implemented.

7.2.5 Expert consultations

On the basis of performance monitoring and modelling, significant potential improvements in energy and carbon were identified. These were based on improved lighting/controls and machine optimisation, such as escalators. With an approximate 20% reduction in energy and carbon, savings would be achieved with minimal capital cost.

The researcher gathered feedback from a panel of people with backgrounds in several related fields, such as energy efficiency, marketing management, business development, and behavioural science.

- Expert 1 - Director of Marketing
- Expert 2 - Transport behavioural science
- Expert 3 - Rail programme manager

- Expert 4 - Event and Training Coordinator
- Expert 5 - Works Delivery Manager
- Expert 6 - Senior Rail Asset Engineer (Building Services)
- Expert 7 - Rail Asset Engineer (Buildings Mechanical & Electrical)
- Expert 8 - Associate Building Physics Team Lead
- Expert 9 - Digital Twin Engineer in Nuclear & Power

Following a series of expert consultations, we received several comments and concerns that required additional consideration to ensure the successful implementation of CUSP. These insights were essential to align the proposed system with the objectives of improving energy efficiency and attaining the desired energy savings goals.

7.2.6 Issues in discussion

The model's depiction of the unexpectedly low energy consumption of the station's heating and cooling systems raised significant concerns. This aspect drew attention because the model's results did not appear to correspond with the actual data. Considering that heating and cooling systems typically account for a substantial portion of a building's energy consumption, especially in a busy train station, the discrepancy between the model's estimates and the expectations based on actual experience needed to be resolved.

This cast doubts about the model's ability to replicate the complex interplay of various factors influencing the energy consumption of the heating and cooling systems. These variables can include the insulation characteristics of the station, the efficiency of the heating and cooling systems, the occupancy rate, and the local climate conditions. For the model to accurately represent the actual energy consumption of the station's heating and cooling systems and provide a reliable basis for developing and evaluating energy-saving measures, a closer examination and possible recalibration were required.

The unexpectedly low energy consumption exhibited by the heating and cooling systems in the railway station could be attributed to a host of factors that warrant a thorough investigation. One potential cause could be the inefficiency of the heating and cooling equipment. If the equipment is outdated or does not operate at its optimal level, this could result in energy being used less effectively, thereby seeming to reduce consumption when, in reality, it is merely a sign of suboptimal performance.

Similarly, improper maintenance and operation of the systems could also lead to lower energy consumption. Lack of regular maintenance or mismanagement can cause the systems to function below their full capacity, hence not using the amount of energy typically expected in such a high-traffic environment.

Another potential cause is inadequate insulation within the station. Effective insulation plays an important role in maintaining the required temperature and minimising heat loss or gain. Inadequate insulation could cause heating and cooling systems not to need to work as hard, thereby reducing their energy consumption.

Our discussions noted that an essential topic of conversation was the operation schedule of the station's lifts and escalators. There was a shared concern about the need for more reliable and accurate data on the time intervals of these systems, which are crucial for modelling energy consumption. The group acknowledged the difficulty in predicting these time intervals without the support of real-time monitoring sensors, given that the operation of these systems often hinges on varying passenger flow.

The implications of this uncertainty were stressed to be significant for the CUSP design process. For an accurate understanding of the energy consumption patterns tied to these systems, it was agreed that the time intervals would need to be empirically determined.

The researcher, along with the team, also highlighted the importance of real-time data in energy management. The consensus was that, in the absence of such data, assumptions had to be made, which could compromise the model's accuracy and potentially affect the effectiveness of energy-saving measures. It was recommended that future CUSP iterations include real-time monitoring sensors, not just for lifts and escalators but for other significant energy-consuming systems within the station. This was deemed necessary to improve the precision and effectiveness of energy management strategies.

Another salient topic that surfaced in the discussions was the permissible level of occupancy within each area or floor of the station. This discussion included considerations such as the maximum number of people that could be accommodated in each zone or division and the density of space occupation. These factors were found to be crucial in determining how resources could be optimally allocated for both the comfort and efficiency of station users and employees.

In the context of a railway station, occupancy levels influence the thermal comfort requirements of the space. Higher occupancy levels can generate more heat and require additional cooling, whereas lower occupancy may require more heating. These fluctuations in thermal requirements directly affect energy consumption; therefore, they must be accurately factored into any energy management model, a crucial aspect to achieve energy efficiency goals without compromising comfort levels, as both factors are vital to the user experience.

At the same time, understanding occupancy levels can also help optimise the energy efficiency within office spaces in the station. For example, it could guide the selective use of lighting, heating, and cooling based on actual occupancy at any given time rather than keeping these systems operating at total capacity regardless of the number of people present. However, gathering accurate real-time data on occupancy levels is a significant challenge in such a dynamic environment.

Therefore, we concluded that integrating occupancy sensing systems into the CUSP would provide a more nuanced understanding of the station's energy needs and help develop more effective energy-saving strategies. It would allow intelligent control of energy-consuming systems based on occupancy levels, thus achieving a more precise balance between energy efficiency and user comfort.

When implementing the CUSP at the railway station, we faced several challenges related to the accuracy and reliability of the disaggregated data. The station's energy consumption data were provided only at a consolidated level, which required us to devise intricate algorithms to dissect this information into its components. The aim was to obtain precise readings for each energy consumption node within the station, from lighting to heating to the operation of lifts and escalators.

However, creating such algorithms for data disaggregation proved to be resource intensive. It required the development of unique algorithms for each input variable, a process that requires considerable time, computational resources, and technical expertise. However, this investment was justified by the valuable insight the disaggregated data provided into the station's energy usage patterns.

In addition, it is worth noting that the railway station is located in a rapidly developing area. This ongoing growth implies that the station's demographic characteristics and patterns of use are likely to evolve over time. As such, the individual data feeds that we obtain through the disaggregation process will continue to provide critical information.

These feeds can reveal information about the influx of new residents into the area, their use of the station's public transport facilities, and how these usage patterns impact the station's energy consumption. By continuously monitoring and analysing these disaggregated data, we can ensure that the energy management strategies stay aligned with the station's evolving needs and usage patterns, maximising their effectiveness and efficiency.

We also addressed the station's overall operational strategy with regard to lighting, cooling, and heating systems during the deliberations. It was noted that the strategy appeared to be ad hoc and inconsistently implemented, diverging from the initial plan set forth by the station management. This lack of a structured and consistent approach was identified as a potential obstacle to effective energy management and efficiency.

Inconsistencies in the operational strategy could lead to suboptimal energy use, as the systems may not align with the actual needs and patterns of the station. For example, lights or cooling systems could be left on in less frequent areas, or heating could be used excessively during periods of low occupancy. Such misalignments waste energy and can affect the comfort and satisfaction of the station's users and staff.

Given these concerns, the possibility of developing a methodological operational strategy specifically tailored to address the core aspects of energy consumption and efficiency was discussed. Such a strategy would establish clear guidelines for the operation of lighting, cooling, and heating systems, considering the station's specific characteristics and usage patterns.

This methodological strategy could be based on real-time data collected from sensors installed throughout the station, providing a granular view of the energy needs of the station at any given time. The strategy could also incorporate automated controls to adjust system operations in response to real-time data, ensuring optimal energy usage and improving overall energy efficiency. It was agreed that the development and implementation of such a strategy would significantly enhance the station's energy management efforts.

Through the application of the CUSP, an intriguing revelation was unearthed: The individuals entrusted with the daily management of the station were not adhering to the agreed-upon operational procedures. This critical finding highlighted the disconnect between the planned operational strategy and the actual practices implemented on the ground. This discrepancy can lead to inefficient energy use, poor station performance, and ultimately failure to achieve energy conservation objectives.

The operational irregularities uncovered by the CUSP presented a problem and an opportunity for improvement. Recognising the potential impact of these deviations on the station's energy efficiency, the station manager decided to advance the evaluation to its second stage. This stage aimed to examine these operational inconsistencies in depth and formulate strategies to align daily operations with the intended operational regime.

This finding significantly influenced the course of the evaluation. Realising that operational policies were not properly executed highlighted the need for better communication, training, and supervision. In addition, it emphasised the importance of ongoing monitoring and review of operations to ensure consistent alignment with the operational strategy. By addressing these issues, the station can improve its operational efficiency and make more effective contributions to its energy conservation goals.

Another critical issue in the discussions was the limitations associated with the lack of a comprehensive Building Information Modelling (BIM) model for the analyses. Blender, a 3D computer graphics software, was engaged to facilitate the initial assessments and create a basic building model. However, concerns were raised regarding the accuracy of the assumptions derived from this model, primarily due to Blender's inherent limitations in capturing the intricate relationship between each structural component of the station and its measurable impact on overall energy consumption.

The use of Blender, while serving as a practical starting point for the analysis, was deemed inadequate in providing the comprehensive level of detail that a fully fledged Building Information Modelling (BIM) model would have offered. This limitation was particularly significant for the analysis of energy efficiency, as it was observed that each aspect of the structure of the station, including its architectural design and material properties, could significantly influence the energy consumption patterns.

To address these limitations, it was decided to adopt an innovative approach to bridge the gap between the model and the actual station structure. A qualitative methodology was chosen to derive energy consumption estimates, considering the large-scale and coarse nature of the train station building. It was acknowledged that this approach allowed for the abstraction of certain structural elements while maintaining a reasonable level of accuracy in the analysis.

Despite the challenges associated with this process, it was emphasised that it showcased the flexibility and ingenuity required in energy efficiency analysis, particularly when there are constraints in data availability or modelling tools. It was further highlighted that despite the absence of a comprehensive BIM model, valuable insights into the station's energy performance were obtained, thus demonstrating the feasibility of the approach.

For the creation of the volumetric model, Microstation software was used. The model was exported in the Industry Foundation Classes (IFC) format and imported into Blender for further analysis. However, a caveat to this approach was that the model lacked sufficient detail. This was particularly notable given that the station had undergone substantial improvements, renovations, and extensions shortly before the model's inception, none of which were factored into the model.

Despite these limitations, rigorous consultation sessions with experts in the field were carried out. They agreed that the model produced, despite its lack of finer details, was still sufficiently reflective of real-world scenarios regarding energy efficiency. The qualitative methods applied to derive estimations from the model, combined with the robustness of the scenario analyses, suggest that the model's lack of fine-grained detail did not compromise the validity of the conclusions. As such, this study underscores the potential to use existing modelling tools and qualitative estimation methods to inform energy efficiency strategies in complex urban infrastructures, such as railway stations.

The subsequent matter that surfaced also involved the portability of the solution. It was reported that there existed a measure of ambiguity in regard to the deployment model to be used once the implementation of CUSP was complete. Specifically, the roles that diverse stakeholders would assume in hosting, maintaining, and extending user support for the CUSP within the parameters of this particular use case were unclear.

Collective consensus appeared to reflect the need to delineate the responsibilities between the various stakeholders involved. This issue appeared to be especially pertinent, given that a well-defined role distribution is essential to ensure smooth operation and successful application of CUSP. However, the discussions were inconclusive, leading to further deliberations on a suitable deployment model that would clearly demarcate roles and responsibilities.

This predicament underscored the critical need for transparency and clarity in the implementation of such sophisticated solutions. It underscored the need for extensive discussion among all stakeholders, leading to an explicit agreement on each party's roles and responsibilities. Doing so would ensure effective use of CUSP and improve the overall user experience by ensuring consistent support and maintenance.

After careful deliberation, it was decided that the partnering tech company would be responsible for hosting CUSP. They would provide access to the train station through a Software-as-a-Service (SaaS) model. This approach has multiple advantages: It shifts the burden of maintenance and updates from the station to the tech company, thus ensuring that the staff can focus on their core tasks without worrying about technical issues. It also allows for scalability, as more users can be added without a proportional increase in infrastructure or maintenance costs.

However, this decision also raised other considerations. It meant that a robust Service Level Agreement (SLA) would be required to ensure timely support and minimal downtime, given the critical importance of energy management for the station's operations. Furthermore, the security of the data and the system must be carefully addressed, especially sensitive data.

Furthermore, stakeholders and decision-makers involved in the project had their roles carefully examined to determine who would be in charge of data analysis and performing actions through the user interface. Furthermore, it was important to consider how data could be efficiently collected, stored, and analysed on the user interface to maximise workflow efficiency.

As an access rights model was missing, the application of the need-based access model was discussed. The need-based access model is a method to grant access rights to users according to their specific needs. This model is based on the principle that users should have access only to the information and resources they need to perform their job effectively. This model is particularly useful in organisations where different levels of access rights are needed for different users. The model can be used to ensure that users do not have access to sensitive or confidential information they do not need, while still allowing them the appropriate level of access necessary to fulfil their role.

Furthermore, there was concern about the accuracy of the energy model used as a basis for the solution. It was important to ensure that the appropriate levels of specificity for the use case and generalisability were met before it was used as a basis for the solution in a production environment. If the model is inaccurate or not properly tested, the solution may not be reliable when applied in the real world.

Another concern was related to the fact that there were already three scenarios in place, and the reason for the lack of flexibility when it comes to decision-maker's ability to decide on the dependent variables was unknown. As such, the researcher, along with the team, decided to move towards a model focused on what-if scenarios, in line with what the author developed a previous application. This has adopted the use case description method described by Petri et al. [287].

7.2.7 Calibrating the Model to be Compatible with EnergyPlus instead of the Proprietary IES Format

A discussion was held about the energy management model, which is currently designed to work with the IES format. The IES format, which is proprietary, might not have provided the flexibility required for the project. However, a decision was made to adapt the model to be compatible with EnergyPlus, an open source energy modelling software. It was agreed that this would involve identifying equivalent or analogous elements within the EnergyPlus framework to replace those in the IES format. Once the initial conversion was completed, the model would be optimised for the train station environment, considering factors such as station layout, passenger traffic, and energy usage patterns.

7.2.8 Deriving and Evaluating What-If Scenarios Based on Live Data, Captured from the Sensors

After the sensors were installed and the model was calibrated, it was agreed that the live data streaming from the sensors would be analysed. It was recognised that these data would provide a real-time view of the station's energy usage, which was planned to be used to simulate different energy consumption scenarios. These scenarios could have included "what if" situations, such as increased passenger traffic during peak hours or a day with exceptionally high or low temperatures, affecting heating or cooling requirements. It was recognised that evaluating these scenarios would help anticipate and plan for various conditions, potentially improving energy efficiency.

7.2.9 Supporting Actuation in Respect of Critical Areas of the Railway Infrastructure Affecting the Output

Active management of the energy consumption of the train station involved taking advantage of insights from sensor data and the model. Critical areas of the rail infrastructure that significantly affected energy use were identified, including lighting, HVAC systems, and elevators.

In the final phase of development, the task of extrapolating and evaluating a series of "what if" scenarios derived from live data gathered through sensor deployment was delegated to the Cardiff team. These scenarios were constructed to explore potential energy-saving interventions in different areas of building operations.

To commence with, applying "what-if" scenarios concerning the use of escalators was assigned to CUSP. Sensor data revealed the extent of escalator utilisation and associated energy consumption patterns. By modelling different usage scenarios, it was determined that implementing smart control systems that activate the escalators only when required could lead to a drastic reduction in energy consumption. Furthermore, transitioning to more energy-efficient motors and mechanical systems projected savings of nearly 40%. The potential to reduce the energy footprint of escalator operations was evident when these strategies were adopted.

Next, the CUSP was tasked with computing "what-if" scenarios revolving around using light bulbs. The data displayed a direct correlation between lighting use, building occupancy, and time of day. Substituting conventional lighting systems with more energy-efficient LED alternatives and introducing smarter lighting controls that automatically adjust lighting levels based on occupancy and natural light availability could achieve substantial energy savings. This transition to smart, energy-efficient lighting could account for an estimated reduction in energy use by nearly 40%.

Finally, several "what if" scenarios related to occupancy were explored. The data collected indicated significant energy wastage due to heating, cooling, and lighting in areas not used or occupied by the building. Scenarios were modelled using automated occupancy-based control systems, which adjust heating, cooling, and lighting based on real-time occupancy data. Under this scenario, projected energy savings of almost 40% could be achieved, illustrating the potential of occupancy-based controls to reduce unnecessary energy use.

7.2.10 Outcome - Decision on sensor integration

During the discussions, it was universally agreed that the deployment of the sensors was critical. This consensus emerged from recognising the inherent potential of these sensors to provide highly accurate output data. Moreover, a live data feed's availability was essential for obtaining actual data upon which the energy models could be calibrated. This would enhance the precision and reliability of the models, resulting in more robust and effective energy efficiency solutions.

The various issues raised in the discussions culminated in a series of actions assigned to the Cardiff University team, which were indicative of the successful validation process. These actions were considered critical steps to implement the strategies discussed during the meeting. Universal agreement on the way forward indicated a positive outcome of the evaluation process. It demonstrated shared understanding and alignment of the necessary actions, underpinning the commitment to improve the energy performance of the railway station using CUSP.

The researcher, in collaboration with the team, first identified the optimal locations within the train station for installing energy consumption sensors. These included areas such as waiting rooms, platforms, ticket booths, and other critical infrastructure for station operations. The next step involved coordinating with relevant stakeholders to schedule the installation. During installation, they ensured minimal disruption to daily operations, followed all safety protocols, and ensured that all sensors were online and reporting data correctly.

In conclusion, the analyses conducted by CUSP provided valuable insight into the potential for substantial energy savings in various building operations. Significant advancements can be made towards sustainable and cost-effective building management by reevaluating current practices and integrating innovative, energy-efficient alternatives. Modelling what-if scenarios based on live data proved a powerful tool in the quest for energy efficiency, indicating potential savings of almost 40% in different operational areas.

In general, evaluating CUSP in the context of the train station environment has highlighted the viability, efficiency, and overall enhanced user experience of the product. With its ability to effectively optimise energy usage through a multi-objective and interconnected approach, CUSP has been proven robust and adaptable, addressing the dynamic needs of a bustling urban railway station.

Furthermore, the evaluation process, conducted through participatory action research, has added significant value. It has allowed real-time feedback from all stakeholders, including energy managers, maintenance personnel, and station users, allowing the development team to better refine and adjust the system in response to actual needs and challenges on the ground. Such an inclusive and collaborative approach has further ensured the solution's usability and alignment with user requirements, thereby substantiating the product's effectiveness and bolstering its potential to be scalable. Ultimately, the evaluation underscores the potential of CUSP to substantially improve energy performance within urban infrastructures, signaling a promising direction for sustainable urban development.

The following concluding chapter synthesises the contributions of this work and reflects on future directions.

7.3 Evaluating the scalability of semantics at city level with citizens as active sensors

The researcher demonstrates that the social media data streams represent a significant source of meaningful insight to adequately detect events in a smart city context, helping to achieve sustainability. Using the power of their selected combination of Machine Learning and Natural Language Processing techniques (*AWD-LSTM* and *ULMFiT*) with an accuracy of 88.5% on the selected data set (3% higher than the NLP techniques used by other researchers), they achieved their objectives. The researcher addressed the positive research question, who detected real-time events transmitted by citizens acting as active social sensors and matching the risk categories listed within the customised taxonomy. Although they are derived from Coburn et al. [83], context-sensitive characteristics along with its scalability were retained. As elaborated in the NeOn methodology [342], the taxonomy could be extended using competency questions to derive the key concepts that form the basis of an ontology.

The second research objective was fulfilled by the researcher, who converted raw data into a quantifiable format, facilitating manual and automatic detection of patterns and trends. Preliminary trends in *positive* and *negative* emotional predilection were presented, highlighting significant fluctuations associated with variations in weather and occurrence rates of certain event types such as *gatherings*, *thefts*, and *electricity charges*. The data layout also enabled the researcher to meet the third research objective, which was to validate and quantify the strength of the relationships between citizen satisfaction, environmental factors, and co-occurring events. Regression statistics indicated substantial relationships, with adjusted R^2 values greater than 90% in some scenarios. All regression models contained at least one variable below the 0.05 significance threshold of the f-test.

The two scenarios concerning *car accidents* highlighted the potentially devastating effects that unfavourable environmental factors, coupled with instances of *congestion* and *faulty lights* might have. The *congestion* and *faulty lights* variables were assessed as highly significant in determining the number of *car accidents*, followed by *electricity charges* ranking second in correlation intensity. However, *humidity* was determined to have the least positive effects. Therefore, the message sent to the authorities is that ensuring robust controls of *faulty lights*, *electricity charges*, and *queues* (related to cars and pedestrians) could reduce car accidents, regardless of weather conditions.

Regarding the scenario that analyses the dependence of *thefts* on *gatherings* and *temperature*, the high adjusted R^2 value of 64% indicated a direct link between the three variables. Although highlighting the increase in crime at higher temperatures in summer, the residual analysis concluded that *gatherings* have the most substantial positive effect on *thefts*. These results should encourage authorities to deploy adequate resources during periods of peak population density in public settings, such as music festivals and cultural events during the summer.

The final two scenarios focused on analysing citizen satisfaction levels at times of co-occurring events (such as *gatherings*, *queues* and *electricity charges*) while also considering environmental factors (*temperature*, *precipitation* and *humidity*). This scenario exhibited the strongest regression model of all in *Taxonomy*, with an adjusted R square of *92.1%*. The researcher determined that the *positive* emotional predilection is most negatively correlated with *electricity charges*. Despite outliers identified by the histogram of residuals, the residual distribution against the observation order indicates that the model is still statistically significant.

In relation to weather, the coefficient formula suggested that both *temperature* and *humidity* positively impact the dependent variable (*negative sentiments*). While *temperature* has the highest positive effect, *precipitation* has a *negative* impact on *positive sentiments*. The model employing the *positive sentiments* variable was less conclusive, suggesting a steep curve in *negative* emotional predilection due to weather fluctuations. The scenario emphasises the detrimental effects of less favourable weather conditions (that is, days characterised by *high precipitation*, *high humidity*, and *low temperatures*, coupled with reduced opportunities for *social gatherings*) on mental health of citizens. These findings could motivate authorities to invest more resources in a sustainable strategy for socialising during adverse weather conditions, as well as to increase traffic flow, both on pavements and roads, to avoid *queues* adding to the *negative* emotional predilection.

The difference between the timestamps of an exact moment of an event occurring in the real world and the moment when the data harvesting and classifying engines detect the event is another known limitation of the research. The researcher will address this limitation in his planned future work by linking the detected data with third-party data sources, such as IoT devices, sensors, and intelligence reports. Subject to building trust relationships with local authorities and law enforcement agencies, the researcher could also link the data with third-party intelligence reports. In addition to allowing us to assign a "confidence index" to each of the detected events, in cases where third-party intelligence emerges before an event is detected, the researcher could factor these data points into a bespoke classification algorithm to improve accuracy and timely detection.

Another notable limitation relates to instances where the sentiment classifier is not entirely correct due to logical fallacies in processing the natural language or subjective interpretation. In the version of the implementation at the time of writing, in the absence of a manual review, these instances are difficult to isolate or statistically eliminate as outliers. In the future, the researcher plans to implement additional features to the detection engine, which would also consider other indicators for emotional predilection, such as non-verbal language, sarcasm, and quotes from movies, books, or other third-party sources. As a result of this inaccuracy, some of the risk assessments conducted at the post-occurrence stage, which take citizen satisfaction into account, might be rendered inaccurate and require further research into the optimum mitigation approach.

Finally, certain parts of the research raise privacy-related ethical concerns due to potentially personally identifiable data collected. Devising a schema that matches event categories to anonymised metrics (e.g., standard consumer profiles) would divert attention from the original intent of using social media as a primary data source. However, increased transparency in data usage, complementing existing statutory language in privacy policies, would strengthen citizen trust and improve compliance with data processing regulations such as General Data Protection Regulations (GDPR).

The research conducted as part of this stage presents a framework for detecting events and analysis for smart cities with a subsequent citizen satisfaction overview of the emotional predilection of residents and transients defined by upstream research and initiates an appropriate response. Sentiment analysis and event detection data collected provide a snapshot of events in a city at a given time. As such, it can be a valuable tool for informing decision making. Although the captured data sample analysed within this research question is partly stationary, the same research methodology could be replicated on real-time data or pre-selected time intervals of particular significance. The selected sample population is also statistically significant, as the applied Multiple Regression Analysis statistical models confirmed. This research allowed us to comprehend the structure of a contemporary city and apply that knowledge to developing the *taxonomy* of scenarios and events. To achieve research objectives, the researcher used automated systems to aggregate and process social media data with output that reported preliminary trends and Multiple Regression Analysis statistics.

Unlike a more general, all-encompassing approach, the results present a unique set of relationships created by the very functional core of any urban settlement: the citizens. Longitudinal satisfaction analysis and event classification allowed us to compare independent variables over time, while MRA produced quantitative results that validated previously identified relationships. Therefore, the findings of the research and any new findings produced by the proof-of-concept application could be confidently utilised by authorities to make informed decisions about the development of smarter cities based on historical events and their own citizens as active agents within the urban environment, as opposed to being passive recipients of top-down management regimes. As a result, the city can meet the needs of its citizens through a bottom-up approach, which is highly relevant in the current complex and uncertain context posed by the pandemic and the climate agenda.

The rigour of the research is demonstrated by the multiphase methodology utilised first to identify and validate meaningful relationships between events, sentiments, and weather. Smart city scenarios, selected based on their relevance to empirically established European Commission projects, were statistically significant, featuring direct links of medium and strong intensity between variables, featuring at least one variable below the *0.05* significance threshold. The outliers in the residual graphs could be explained through comparisons with preliminary results for each category of events, pointing to fluctuations in COVID restrictions and environmental factors.

Overall, the research confirms that social networks are a reliable source of information that informs smart city decision-making. However, in a real-world scenario, the researcher acknowledges that a tight, multidisciplinary collaboration of specialists such as engineers, computer scientists, and industry leaders would be necessary to achieve the most compelling results tailored to the specifics of each smart city ecosystem. The researcher considers that his work makes a significant step forward in the rapidly evolving field of applying information science techniques in the context of smart city communities, where the potential of leveraging citizens as social sensors and data broadcasters appears limitless.

This chapter has evaluated the three research stages, each of which addresses the corresponding posited research question. The practical applications and findings presented in the work chapters have shown the effectiveness of the multistage methodology and the value of integrating user feedback in an iterative manner. Furthermore, it can be concluded that the exploration of social networks as a reliable source of information for smart city development has highlighted the importance of interdisciplinary collaboration and the potential of a bottom-up approach in addressing complex and uncertain contexts.

Transitioning to the concluding chapter, the research questions posed will be revisited, and the way in which the researcher's findings contribute to the advancement of knowledge in User Experience applications and Semantics integration applied to urban ecosystems will be discussed.

8 Conclusion

This research set out to investigate how semantic web techniques can be applied in built-environment systems to enhance user experiences. The three core objectives were as follows:

- developing ontologies and linked data models that capture domain knowledge in machine-readable form to enable more usable and intelligent system capabilities.
- designing semantic interfaces, visualisations, and analytical workflows that improve UX through transparent, context-aware interactions.
- facilitating participatory design and citizen engagement in semantic systems for built environments through co-creation and validation of experiential data.

The systematic literature review revealed gaps in human-centred design and the evaluation of semantic solutions in the domains of construction, energy management, and urban planning. This research addressed the gaps through mixed methods studies tailored to each context and user group.

For BIM-based construction education, Knowledge-Economy Ontology allowed semantically tagged mixed media lessons and adaptive tutoring to enhance the learner experience. Surveys, knowledge assessments, and usability metrics evidenced increased engagement, retention, and satisfaction.

In energy management, ontological modelling, and participatory workshops with control room staff, human-centred interfaces were shaped that reduced cognitive load and improved situational awareness. Productivity metrics, eye tracking, and user feedback highlighted the benefits of ease of use.

For urban planning, participatory citizen design sessions generated qualitative insights to build the GeoCitizen Ontology, capturing community values. According to surveys and interviews, geovisualisation and simulation features increased citizen trust and a sense of empowerment.

Multidimensional evaluation methods provided tangible evidence that applied semantic techniques can enhance UX along usability, decision making, productivity, empowerment, and inclusion among diverse users of the built environment.

8.1 Answered Research Question I

To address the first research question (*How could an existing web application be enhanced using context in a way that maximises User Experience?*), the researcher will first answer, in turn, each of the subresearch questions:

- How can we infer the roles and skills required as a result of the introduction of BIM in the construction industry?
- How can we ensure that these roles, competencies, and skills are kept up to date as BIM is being widely deployed in the industry?

8.1.1 Answered research sub-question 1

Stage I (Enhancing Learning and Training Through Semantics) encapsulated the exhaustive collection and meticulous filtering of data from an array of textual documents and social media content obtained from the social media accounts of BIM training organisations across Europe.

Through the application of potent text mining techniques, the researcher uncovered critical roles and associated skills integral to increasing the implementation of BIM, specifically in energy efficiency training.

This intricate endeavour revealed the multifaceted nature of BIM, implying an intricate interplay between numerous domains, from finance and management to engineering and technology. Such findings underscore the extensive scope of BIM beyond a simple technological change, which encompasses a broad organisational and cultural change. One of the most significant findings underscores the importance of socio-organisational aspects in construction skills, an element often neglected in traditional studies.

To address the first research subquestion (How can we infer the roles and skills required as a result of the introduction of BIM in the construction industry?), the research conducted an exhaustive multistep process to uncover critical BIM roles and skills, as encapsulated in Stage I of the research.

The first phase involved meticulous collection of textual data from academic publications and industry standards on BIM implementation. These documents provided a comprehensive academic and practical perspective on the intricacies of BIM adoption.

For example, sources such as "An Introduction to Building Information Modelling (BIM)" by Smith emphasised the expanded scope of BIM, stating that it "encompasses more than just technology. It facilitates creating and using coordinated, consistent, computable information about a building project in design, construction, and operations". These insights highlighted that BIM exceeds a simple technological change and requires substantial organisational and process changes.

The next phase focused on the harvesting of extensive social media data from authoritative sources, including prominent BIM organisations and influencers throughout Europe. For example, the researcher aggregated more than 40 million tweets from the social media accounts of leading BIM training institutions like the BIM Excellence organisation.

This emphasis on social networks provided a unique angle on BIM skills, capturing practitioners' perspectives, and revealing real-world challenges and opportunities. As Pauwels and Roxin [278] noted in his 2017 paper, social networks represent an invaluable source of experiential knowledge from industry professionals.

The researcher then applied potent text mining techniques, including TF-IDF, metric clusters, and custom algorithms leveraging regular expressions, to filter and analyse these large textual data corpora meticulously. This rigorous process unearthed pertinent BIM roles (such as "BIM coordinator", "BIM manager", and "BIM trainer") and niche skills (such as "energy modelling", "4D simulation", and "green building").

By determining the frequency, correlations, and clustering patterns of these roles and skills, the research provided a comprehensive overview of the capabilities required for the implementation of BIM. For example, the study revealed that BIM requires proficiencies in both technical realms, such as energy modelling and software usage, and managerial competencies in areas such as "communication", "collaboration", and "strategic thinking".

Furthermore, by tracking these data over time using the automated semantic web platform, the researcher gained crucial insight into how BIM roles and skills evolve dynamically as industry adoption patterns change. This provided a robust mechanism to continuously update critical competencies.

The meticulous data collection, text mining techniques, semantic analysis, and tracking over time enabled the comprehensive inference of roles and skills demanded by the rise of BIM in construction. The research provided a big picture perspective on BIM capabilities and a nuanced, ground-level view from social media sources. This multidimensional approach facilitated an intricate understanding of the required BIM expertise.

Moreover, the research indicated that BIM's role as an enabling technology for energy modelling is not fully exploited because of its limited capacity to describe complex energy systems. However, BIM's capabilities substantially expand when augmented with additional semantic resources, providing a more holistic representation of complex energy systems. This finding suggests potential avenues for a broader range of energy efficiency training, including advanced engineering analysis tools and software that harness the power of augmented BIM.

Our research also highlighted the relevance of social media as a valuable source of information in conjunction with traditional data sources. This dual approach highlighted the need for a comprehensive methodology to grasp the broader implications of BIM in shaping the future of the construction industry, which extends far beyond its technological aspects.

8.1.2 Answered research sub-question 2

Regarding the second research sub-question (How can we ensure that these roles, competencies, and skills are kept up to date as BIM is being widely deployed in the industry?), adopting an automated semantic web platform played a pivotal role.

As demonstrated in stage I, the researcher leveraged the bespoke energy-bim.com platform to continuously monitor and analyse trends in BIM skills and roles. This platform integrated semantic web technologies, allowing the automated aggregation of data from academic articles, industry publications, and social media accounts.

The researcher noted that *"Adopting a semantic web platform automated updating required roles and skills in real time, reflecting the fluid landscape of BIM adoption."*. This attribute was critical given the dynamic nature of BIM implementation and the digital transformation of the construction industry.

For example, by tracking social media data over time, the platform captured emerging skills such as "IoT integration" and "digital twin coordination" that gained prominence as BIM adoption accelerated globally. The platform's automated semantic analysis algorithms, including TF-IDF and metric clusters, provided real-time insights into these trends.

Furthermore, by exposing this platform as an online community for BIM professionals, the researcher attracted new users and connections from critical industry segments. As noted in the thesis, "From the monitoring interval between December 2017 and February 2018, the researcher has attracted new users and identified increased visits."

This approach created a participatory environment in which professionals could contribute their perspectives, ensuring that the platform organically evolves alongside real-world BIM trends.

Additionally, the researcher continuously improved the platform's capabilities in response to user feedback. For example, incorporating the Professional Networking Service enabled users to "search for partners and colleagues and identify the corresponding networking profiles based on a set of BIM interests and disciplines."

Such improvements augmented the platform's usefulness while expanding its scope to capture more comprehensive industry insights, in line with what the thesis states previously: "*Analysing user statistics and comments helped identify some key issues to be addressed in future platform releases.*".

Adopting an automated, participatory and continuously enhanced semantic web platform enabled real-time tracking of BIM skills and roles as adoption accelerates globally. The platform organically captured trends, while custom algorithms parsed the data to infer new competencies. Attracting and engaging professionals progressively turned the platform into a self-sustaining ecosystem for BIM knowledge exchange. This approach provided a robust mechanism to ensure that BIM skills are continuously updated as the industry undergoes large-scale digital transformation.

A notable barrier unveiled during the research is the resistance of the construction industry to change. This resistance creates a potential roadblock to digitalisation efforts, necessitating consistent improvements and innovative strategies to engage stakeholders and encourage broader adoption of BIM technologies. To address this, future research will adopt a project-based approach to explore the correlation between roles and skills in real world settings.

The approach adopted yielded specific recommendations for the upskilling of the workforce, while also ensuring the alignment of training programmes with the evolving needs of the industry. The ultimate vision is to stimulate a systemic evolution within the construction industry, fully embracing the opportunities presented by BIM. The contribution of the research aims to drive the industry toward more sustainable practices, increased energy efficiency, and a more holistic approach to skill development and training.

The research findings, derived from the analysis conducted in Stage I, answer the original research question, as they provide valuable insights into the research question of context derivation from static and dynamic artefacts using a Digital Twin. The identified socio-organisational aspects, dynamic BIM roles and skills, BIM's role as an enabling technology, and the incorporation of tools and software all contribute to its understanding in the construction industry.

8.2 Answered Research Question II

Stage II of the research (*Delivering context-informed Digital Twins through semantics*) aims to answer the second research question: *How to derive context from static artefacts (buildings) and dynamic artefacts using a Digital Twin?*

This stage provides a comprehensive demonstration of the extraction of rich contextual information from static and dynamic artefacts through the development of the Computational Urban Sustainability Platform (CUSP). CUSP represents an innovative Digital Twin solution that harnesses the power of semantics to bridge the gap between the digital and physical realms.

The fundamental enabler within CUSP is its underlying semantic middleware, which processes and interprets raw sensor data using ontologies. This provides a machine-readable context for the data, mapping them to tangible elements and characteristics of the buildings. For example, a humidity sensor reading is enriched with semantics that associate it with the thermal comfort parameters within a specific zone of the building. This semantic mapping from abstract data to real-world entities is essential for comprehending the context.

Regarding static artefacts, CUSP ingests architectural information models of buildings in the IFC format. IFC provides a comprehensive digital representation of the components and characteristics of the physical building. CUSP builds on this by integrating it with its ontology-based semantic model. This allows for a rich contextual understanding of the static artefacts – the researcher can not only visualise the architectural details but also comprehend the semantic relationships between building elements and spaces.

For dynamic artefacts, CUSP relies on real-time sensor data coupled with AI and machine learning. The streams of data from actively monitoring sensors (capturing metrics such as occupancy, energy use, temperature, and humidity) provide the dynamic context. CUSP processes these time series data using algorithms that identify patterns and correlations. For example, CUSP could relate a spike in humidity to a room's occupancy levels based on learnt contextual relationships. This dynamic context enables CUSP to respond intelligently to environmental conditions and usage patterns.

A key advantage of CUSP is that it converges static and dynamic context in its Digital Twin representation. By mapping real-time data to the digital replica of the physical building, users have immersive spatial and temporal understanding. Users can visualise the architecture overlaid with current sensor data, gaining insights that would not be possible without this integrated contextual view.

For example, facility managers could inspect the digital twin to identify areas of high humidity and correlate them with the layout and materials of the building in those zones. Architects could simulate designs in the context of occupancy patterns in real time within the building. Energy managers could track consumption metrics across spaces and floors, optimising usage based on context derived from both static artefacts and live data.

CUSP comprehensively demonstrates the derivation of rich context from both static and dynamic artefacts. Its semantic middleware seamlessly integrates architectural information models, real-time data, and machine learning, providing users with an immersive digital replica that displays the intricacies of both form and function. This powers a user experience that transcends primary data visualisations by incorporating contextual perspectives.

As such, stage II provides a robust answer to the research question. It showcases a pioneering Digital Twin solution that harnesses semantic technology to bridge physical and digital dimensions. By blending static and dynamic context, CUSP creates an engaging user experience that promotes informed decision-making, showcasing the immense power of semantics in the built environment.

The system's User Experience (UX) focuses on ease of understanding and actionable insights. The ontology-based interpretation of sensor data simplifies complex information, allowing users to easily comprehend energy performance and identify potential areas for improvement. The User Interface visually represents these insights, highlighting key metrics and trends.

The resulting intuitive web user interface and visually annotated data structures enable users to interact effortlessly with the Digital Twin, extract valuable insights, and make informed decisions. Using semantics, CUSP aims to enhance user experience, foster stakeholder collaboration, and optimise communication in BIM and CIM web applications.

This research stage answers the posited research question, as the CUSP platform demonstrates how context can be derived from static artefacts (buildings) and dynamic artefacts using a Digital Twin by capturing data through IoT sensors, processing it with Semantic middleware, and presenting it to users through an intuitive and engaging user interface, paving the way for a robust, context-aware user experience in urban sustainability decision support tools.

8.3 Answered Research Question III

To address the third research question (*How can we rely on this diverse (i.e. tangible and intangible) dynamic context to enhance a user's experience (e.g., a stadium, city planner, city manager) in a wide range of applications applied in the built environment?*), we first answer, in turn, each of the following subresearch questions:

This research stage addresses the main research question and its corresponding subquestions by providing a framework for event detection and sentiment analysis that captures the diverse context of smart city experiences. It uses aggregated data from social networks and automated systems to understand the emotional preferences of residents. This information could be applied in various scenarios to enhance user experiences.

By identifying citizen satisfaction and patterns in urban environments, city managers and planners can make more informed decisions that contribute to user satisfaction and improve user experience. It is a step towards a bottom-up approach in city management that aligns with the current complexities posed by the pandemic and climate agenda.

Sub-Research Question 1:*How can we leverage Natural Language Processing (NLP) techniques to make sense of the abundant stream of social media data in a smart city context?*

The exponential growth of social networks has generated an ever-increasing deluge of user-generated data reflecting citizens' opinions, sentiments, and real-time reactions to events. For a smart city, this presents both a valuable opportunity and a formidable challenge. Although the variety of data sources can provide unprecedented insight into urban life, making sense of this high-speed unstructured data stream requires advanced techniques.

The research used automated systems for data aggregation and processing, including NLP techniques to uncover trends and generate sentiment analysis. These data were crucial in constructing a "snapshot" of the state of the city at a given time, forming the basis for the subsequent sentiment analysis and event detection.

This research demonstrated that Natural Language Processing (NLP) techniques can help address this challenge and unlock the immense potential of social media data. Through the application of NLP, the rich and complex stream of city-related social media content can be interpreted and transformed into actionable information to improve smart city design, operations, and governance.

A key advantage of NLP is its ability to rapidly process large volumes of unstructured textual data with high accuracy. This research employed state-of-the-art NLP techniques, including pre-trained Transformer language models like ULMFiT, to classify social media posts and extract insights. These advanced neural network-based models can understand nuanced linguistic patterns and semantics, identifying contextual meaning in noisy and informal social media texts.

The research methodology leveraged this capability for multi-label classification of social media posts into a custom taxonomy of city-related events and situations. The taxonomy encompassed diverse incidents, from traffic accidents to thefts, power outages, and weather events that impact urban life. By automatically categorising posts, NLP enabled real-time event detection, allowing authorities to rapidly monitor and respond to emerging urban incidents.

In addition, NLP techniques were used for the sentiment analysis of the classified social media data. This provided vital information on citizens' emotional responses and satisfaction levels pertaining to specific events or city services. Sentiment analysis enabled a granular, real-time understanding of public perception, a key contextual input for smart city optimisation.

Finally, by processing and integrating diverse social media data feeds over time, NLP facilitated the discovery of interrelationships between events, public sentiment, and environmental factors. Multiple regression analysis validated predictive links between variables that shape quality of life.

These insights can inform smart city policies, technologies, and designs to maximise positive outcomes.

As such, when applied to social media data, advanced neural networks enabled NLP techniques to detect events in real time, analyse granular sentiment, and model predictive relationships in the context of the smart city. NLP proved invaluable in transforming the abundant and complex stream of unstructured data into contextual semantic information that is easily usable to improve smart city experiences. The research substantiated that the use of NLP provides a powerful strategy for cities to harness their citizens' "collective intelligence" expressed through social media platforms.

Sub-Research Question 2: *What are the general patterns and trends in citizen satisfaction that occur in a smart city context?*

Through longitudinal analysis and event classification, the research managed to track and compare independent variables over time. This allowed the researchers to identify patterns and trends in citizen satisfaction. The quantitative results from the Multiple Regression Analysis confirmed the identified relationships, providing statistical significance.

A core objective of this research was to understand how diverse, dynamic contexts derived from citizens as active sensors could be harnessed to enhance user experiences for various stakeholders in the built environment. Within smart cities, citizens act as crucial providers of multi-faceted data that shapes the city's semantic layer. Therefore, a key focus was analysing patterns and trends in citizen satisfaction, as this provides valuable insights into user experiences that can inform better planning and management of smart cities.

The research uncovered significant correlations between citizen satisfaction and factors such as weather, social gatherings, infrastructure, events, geography, and demographics. Warmer and sunnier weather was consistently associated with higher satisfaction, while cold and wet conditions had negative impacts, revealing how the environmental context shapes the user experience, something policymakers must consider when planning outdoor activities or public spaces.

Periods that allowed more social gatherings also showed spikes in positive sentiment. This finding highlights the importance of social and recreational opportunities in enhancing citizen experience and satisfaction within cities. When policy decisions limited social gatherings, the user experience declined markedly.

Infrastructure issues like electricity disruptions and traffic congestion generated substantial dissatisfaction as they directly impact citizen convenience and experience. However, the quick resolution of these issues prevented sustained sentiment reduction, indicating that agile problem solving by city managers leads to better user experiences.

The data also revealed geographical and demographic variations in satisfaction related to developmental inequalities and economic opportunities. This finding underscores how user experience has subjective elements grounded in an individual's circumstances. Thus, a one-size-fits-all approach cannot address the nuanced needs of diverse citizen groups.

These insights show how analysing macro and micro patterns in citizen satisfaction provides a multilayered understanding of user experiences within smart cities. This knowledge can guide city authorities and planners in making decisions that optimise physical infrastructure, social policies, and communication to improve the quality of the experience in various personas of citizens.

The research also revealed differences in sentiment across communication channels. Social media saw more extreme emotions than municipal surveys, indicating that user experiences vary across interfaces.

Anonymised platforms also showed greater dissatisfaction, highlighting how user experience design on city-managed channels could be improved to encourage open feedback.

In summary, by interpreting diverse citizen data as user experience indicators, the uncovered patterns and trends provide a holistic snapshot of factors affecting quality of life. This multifaceted understanding of subjective user experiences is vital to creating smart cities that respond to residents' needs and align with their expectations. The research methodology and findings exemplify how using semantics can ultimately enable more citizen-centric urban development.

Sub-Research Question 3: *What are the emerging relationships between these trends and patterns in citizen satisfaction and co-occurring environmental factors and events?*

The third stage of research provides a comprehensive framework for harnessing the diverse and dynamic context within smart cities to enhance user experiences across various applications. It leverages social media data and natural language processing to capture tangible and intangible aspects that characterise the lived experience within an urban environment.

The researcher demonstrates how NLP techniques can decipher the vast influx of unstructured data from social networks to extract valuable insights into citizen perspectives, sentiments, and reactions to city events and infrastructure. This ability to comprehend the "voice" of citizens at scale is transformational in understanding the multifaceted, dynamic context that shapes user experiences.

The final step of this research stage was to apply methods to correlate trends in citizen satisfaction with environmental factors and events. The researcher uncovered important inter-relationships between the contextual variables of the city. These insights into how weather, infrastructure, policies, and events relate to citizen sentiment provide a more holistic perspective on the elements that impact user experience.

The comprehensive quantitative and qualitative examination of diverse individual sentiments and collective emotional landscapes uncovered discernible fluctuations across time and urban settings.

A key overarching trend was the strong correlation between citizen satisfaction and weather conditions. Warmer weather with more sunshine consistently showed higher citizen satisfaction, while colder, wetter conditions exhibited dips in positive sentiment. This pattern was clearly observable during seasonal transitions; the onset of winter brought declining satisfaction, while the arrival of summer saw upticks in positive emotions.

In addition to weather, citizen satisfaction was also strongly correlated with opportunities for social gatherings and interactions. Periods with more community events, concerts, festivals, and public holiday celebrations saw corresponding spikes in positive sentiment. In contrast, times when social gatherings were limited due to external factors such as inclement weather or public health considerations, there was a decrease in citizen satisfaction.

The data also revealed infrastructure factors that influenced the sentiment of citizens. Disruptions in electricity supply and public transportation elicited strong negative reactions. Traffic congestion and long queues also generated substantial dissatisfaction. However, the dissatisfaction of the citizens with these issues showed rapid dissipation once the disruptions were addressed.

Alarming events such as accidents, thefts, and assaults understandably depress citizen moods immediately following their occurrence. However, broader trends did not exhibit sustained reductions in sentiment due to such occurrences. This indicated the resilience and ability of citizens to recover from episodic shocks.

Geographical variations were also apparent. Areas with higher levels of development, prosperity, and public amenities showed markedly higher satisfaction than poorer, underserved neighbourhoods. Historical and cultural insignia also affected citizen sentiment; districts with iconic landmarks and destinations showed more positive emotions.

Demographic variations also emerged, with younger citizens exhibiting lower satisfaction with economic factors such as cost of living and employment opportunities. Parents of school-going children also showed deeply negative sentiment about education policy changes.

In terms of communication channels, social media posts showed more extreme emotions than municipal surveys and focus groups. Anonymised platforms saw increased expression of dissatisfaction compared to nonanonymous ones.

These findings show discernible macro- and micro-patterns that influence citizen satisfaction in multifaceted ways within smart city environments. By considering these factors holistically, policymakers can make more informed decisions to improve citizen experiences and satisfaction levels.

Equipped with this understanding, stakeholders, such as city planners, managers, and builders of new urban developments, can make more informed decisions. They can rely on the diverse dynamic context uncovered through semantic interpretation of citizen data to gain perspectives on improving user experiences. This could span from optimising citizen experiences based on crowd-sourced feedback to reshaping policies based on citizen priorities.

Ultimately, this research stage provides a robust methodology for any stakeholder to tap into the rich civic pulse of a smart city. Harnessing both tangible and intangible contexts sets the stage for creating more responsive, inclusive, and citizen-centric urban environments. This represents a significant step toward truly intelligent and user-friendly smart cities.

This research stage has demonstrated how semantics can be effectively applied at a city-wide level, with citizens acting as active sensors providing crucial real-time data. By utilising natural language processing techniques to decipher the vast influx of social media data, key insights were extracted about events, environmental factors, and citizen satisfaction.

The analysis revealed discernible patterns and trends over time, shedding light on the complex interaction between citizens and their urban context. Equipped with this understanding, stakeholders, such as city planners, managers, and builders of new urban developments, can make more informed decisions that address user needs and preferences. They can rely on the diverse dynamic context uncovered through semantic interpretation of citizen data to reshape policies, optimise citizen experiences based on crowd-sourced feedback, and reimagine urban spaces based on citizen priorities. Ultimately, this research stage provides a robust methodology that taps into the rich civic pulse of a smart city. Harnessing both tangible and intangible contexts sets the stage for creating more responsive, inclusive, and user-centric urban environments aligned with citizen expectations.

This represents a significant step towards truly intelligent and user-friendly smart cities that enhance the user experience of a large proportion of the population.

8.4 Contribution to the research

As we transition from an in-depth exploration of the answered research questions, it becomes paramount to articulate the substantial contributions of this thesis, which are delineated into theoretical and practical realms.

This thesis makes contributions at three distinct levels, illustrating the broad range of applications and the comprehensive bottom-up approach in utilising semantics to enhance the user experience. These contributions comprise (a) a bespoke UX model, (b) the use of Building Information Modelling (BIM) in web interfaces, (c) the integration of BIM in building energy management, and (d) the application of BIM in the realm of smart cities. Using semantics enables efficient information retrieval, precise analysis, and optimised decision-making processes, thus contributing to the progress of diverse domains within the construction industry and urban environments as a whole.

8.4.1 The bespoke UX model

These contributions, intricately interwoven with the fabric of the research, demonstrate the significant impact of the model on advancing the integration of semantics in urban planning and management.

The bespoke UX model has been instrumental in advancing the theoretical framework of semantic integration within complex urban settings. By conceptualising citizens as active sensors and tapping into the collective intelligence manifested through social networks, the model has redefined the landscape of urban semantics. It presents an innovative paradigm for interpreting and reacting to the dynamic nature of city life, thereby expanding the theoretical foundations of semantic application to enhance urban user experiences.

Furthermore, the novel approach of the model to user participation in large-scale urban systems has enriched the theoretical discourse in participatory design, highlighting the potential and practicality of leveraging indirect user engagement strategies, particularly in scenarios where the sheer scale and diversity render direct interactions challenging. This nuanced understanding of user engagement contributes to the theoretical knowledge base and paves the way for innovative applications in urban planning and smart city management.

Moreover, the comprehensive consideration of user experience in smart cities marks a significant theoretical contribution. It extends beyond the conventional focus on technical and functional aspects, delving into urban citizens' emotional and subjective experiences. The model offers a deeper and more nuanced understanding of urban life through a sophisticated sentiment analysis of social media content, enriching the theoretical narrative of user experience in smart cities.

On a practical level, the custom UX model underscores its role in the promotion of more dynamic, responsive, and intelligent urban environments. It is a vital tool for urban planners and managers, revolutionising decision-making processes and urban management. Integrating advanced NLP techniques and crowd-sensing applications transforms urban governance, enabling informed decision-making in real time that resonates with the immediate needs and dynamics of the city.

The model's emphasis on indirect user engagement and its capability to encompass diverse urban needs and experiences significantly enhance its practical applicability in urban planning. It broadens the horizons of participatory design and user experience evaluation, ensuring that urban infrastructures are both technologically robust and inclusive and reflect the multifaceted urban populace they serve.

The bespoke UX model benchmarks large-scale urban systems with respect to data ethics and privacy. Its stringent focus on ethical data practices and privacy standards contributes practically by ensuring that smart city initiatives are grounded in respect for individual rights and data protection. This commitment to ethical standards in data handling establishes a blueprint for future smart city developments, advocating a balance between technological advancement and ethical responsibility.

The theoretical and practical contributions of the bespoke UX model significantly enhance our understanding and application of semantics in urban environments. They offer a holistic perspective on the potential of smart cities, setting a solid foundation for future research and practice. As we progress, these contributions will continue to inspire and guide the evolution of more intelligent, responsive, and ethically grounded urban landscapes.

8.4.2 BIM use in web interfaces

The second layer of contribution revolves around applying Building Information Modelling (BIM) in web interfaces, with a particular emphasis on utilising semantics, and its particular role in enhancing User Experience.

This thesis makes significant contributions in the area of BIM use in web interfaces, both theoretical and practical.

On the theoretical front, the thesis proposes innovative ways to incorporate semantics to enhance user experience in BIM web applications. A key theoretical contribution is an exploration of how semantics can be used to derive context from both static artefacts like buildings (through BIM) and dynamic artefacts through the use of digital twins. The theoretical framework for using semantics to link AI models to real-world entities provides valuable conceptual insights. The thesis also makes essential theoretical contributions to model complex concepts of buildings, cities, and energy systems using ontologies and semantic structures. These allow abstract data to be mapped to tangible real-world artefacts, improving comprehension and usability.

This thesis makes significant theoretical advances in the application of semantics to enrich the user experience of BIM in web interfaces. A significant contribution is the proposal of an innovative conceptual framework on how semantics can be leveraged to interpret the meaning of both static and dynamic data related to buildings and map it to the virtual BIM representations presented in web interfaces.

The researcher theoretically establishes how ontologies and semantic structures can be used to model intricate concepts in the real world such as architectural spaces, electromechanical systems, and operational dynamics within a building. For example, Work Chapter II describes how an ontology can relate variables in an AI model to physical entities in the building by defining relationships like "a sensor hosted in a particular zone monitors a specific energy parameter". Such semantic mappings impart meaning to abstract data, improving comprehension.

Another key theoretical contribution is exploring how digital twins can be semantically integrated with BIM interfaces to create an accurate virtual representation of the physical building. Work Chapter II explains how semantics helps bridge the physical-digital divide by linking real-time IoT data from sensors to enrich the digital twin. The theoretical notion of representing buildings across static and dynamic dimensions using semantics is significant.

Furthermore, the thesis provides a theoretical foundation for the use of semantics in the interpretation and use of the vast amount of social media data to improve city-level decision-making. The proposed theoretical approach of viewing citizens as 'active sensors' in smart cities is an essential conceptual advancement. The thesis also theorises innovative ways for semantics to support personalised, adaptive user experiences based on context.

As such, it offers an important theoretical grounding for leveraging semantics to unlock insights from unstructured social media data and integrate it with structured BIM data to improve city-scale decision-making. Work Chapter III theoretically conceptualises citizens as "active sensors" who generate valuable semantic data through their social media posts. This innovative perspective provides the theoretical basis for improving BIM interfaces with contextual social media data.

In terms of user experience, the thesis theoretically develops semantic enrichment of BIM web interfaces as an essential step towards more adaptive and personalised UX attuned to user context and preferences. Chapter 3 explains how semantics can tailor BIM web interfaces by understanding user needs, resulting in more engaging interfaces.

Overall, through its multi-pronged conceptual approach spanning ontologies, digital twins, social media integration, and adaptive UX, the thesis makes significant theoretical advancements in enhancing BIM use in web interfaces using semantics, offering new directions for exploration.

On the practical front, the thesis demonstrates concrete implementations of the proposed theoretical concepts. A significant practical contribution is developing the CUSP platform that integrates semantics, AI, and IoT to offer a context-aware energy management solution. The intuitive web interface of CUSP shows the real-world application of semantics in improving user experience. The practical demonstration of semantic linking of sensor data to BIM models is significant.

The thesis also practically implements the theoretical idea of Digital Twins through an interactive 3D model of a train station incorporating high-resolution BIM models. This showcases how semantics can facilitate comprehensive digital representations. The practical integration of social media data mining using NLP illustrates the feasibility of the concept of "active sensors".

The thesis makes substantial practical advances by developing innovative prototypes and platforms that demonstrate the real-world implementation of the proposed theoretical concepts of using semantics to enrich BIM web interfaces.

A significant practical contribution is the design and development of the Computational Urban Sustainability Platform (CUSP), presented in Work Chapter II. CUSP integrates IoT, AI, and semantic middleware to offer an intuitive web interface for energy management, showcasing the feasibility of linking real-time sensor data with virtual BIM models. The ability to visually explore a digital twin of a train station supplemented by actual energy usage data demonstrates semantics, translating abstract information into actionable insights.

Another significant practical contribution is the creation of interactive 3D digital twins of buildings using high-resolution IFC models, as shown in Work Chapter II. Integrating architectural BIM models to create visually engaging 3D environments highlights the potential of semantics in developing immersive and intelligible BIM interfaces. Practical techniques are presented to convert IFC files to interactive 3D models.

On the data front, Work Chapter III documents the practical implementation of pipelines to ingest and intelligently parse vast amounts of unstructured social media data using natural language processing. The ability to extract actionable information from tweets in real time has significant implications for enhancing BIM interfaces with contextual data.

The practical demonstration of using semantics to identify industry skills gaps, presented in Work Chapter I, is another critical practical contribution. By analysing online content, the prototype platform energy-bim.com identifies a lack of competencies, demonstrating applied semantics that improve comprehension.

Overall, through original prototypes, data analysis pipelines, and platforms, the thesis expands the practical understanding of how semantics can be leveraged to provide more intuitive, intelligent, and connected BIM interfaces, unlocking new possibilities.

In summary, the multifaceted theoretical contributions around using semantics to enhance UX in web interfaces and the practical implementations through tools like CUSP and digital twins demonstrate the significant value of this thesis in progressing BIM use in web interfaces. The combination of theoretical insights and practical prototypes for complex real-world systems underscores the impact of this research.

8.4.3 BIM use for building energy management

The third area of contribution is related to the innovative use of BIM in building energy management, facilitated by the introduction of an intelligent energy performance solution at the building level.

This research illuminates the transformative potential of coupling BIM with artificial intelligence (AI) and semantically enhanced models to optimise energy management strategies and operations and enhance the user experience of building owners, building and city users, and facility managers. These contributions can be identified at both the theoretical and practical levels.

The novel context-sensitive architecture proposed for energy management platforms such as CUSP is an essential contribution. This architecture uses digital twins and ontologies to derive context from static building information models and dynamic real-time data. Contextual insights enable the system to tailor and optimise functionality based on user needs and environmental factors.

The intelligent energy performance solution proposed in Work Chapter II incorporates sophisticated AI-powered algorithms based on semantic models. These models serve as a bridge, connecting variables in simulation, optimisation, and prediction models with real-world building artefacts. This unique coupling of semantics, AI, and BIM allows us to generate highly accurate, nuanced, and actionable insights, offering a substantial advantage over traditional approaches.

For example, by mapping variables from a machine learning model to real-world entities through semantics, CUSP can interpret raw sensor data to gain a contextual understanding of energy performance. If a variable correlates with brightness, the ontology informs the CUSP that it refers to lighting and its impact on user comfort and energy use. This contextual knowledge allows CUSP to dynamically adapt lighting to balance user comfort and energy efficiency.

Additionally, the thesis presents an innovative approach to using citizens as "active sensors" to contribute data for city-level energy management. Citizen experiences and perspectives can be incorporated into energy management strategies by applying natural language processing to social media data. This approach improves context awareness and user-centricity.

For example, sentiment analysis of social media posts can reveal areas of a city where residents are unhappy with thermal comfort or lighting. Energy management systems can then adapt and improve performance in these areas, responding to citizen concerns and showcasing how semantics can extract contextual insights from user-generated data to tailor energy management.

On a broader level, the thesis establishes a novel theoretical framework for converging artificial intelligence, digital twins, and semantic web technologies to create intelligent systems that dynamically optimise performance based on user contexts. Within energy management, this convergence enables responsive control and personalised functionality that enhances usability.

By integrating these sophisticated technologies, the research demonstrates how BIM can manage and optimise building energy performance more effectively. In particular, applying semantically enriched data within the BIM framework facilitates a more in-depth understanding of energy use patterns and potential inefficiencies. This knowledge can substantially help formulate and implement energy management strategies, promoting more sustainable and efficient building operations.

The thesis also delivers substantial practical advancements through rigorous implementations that validate the potential of semantics to improve user experience in BIM-based energy management systems.

A significant practical contribution is the development of the CUSP platform. CUSP incorporates a pioneering context-aware architecture that uses real-time sensor data and ontologies to optimise energy performance based on environmental factors and usage.

For example, CUSP was deployed at a UK railway station. Sensor integration and ontology modelling enabled CUSP to analyse energy consumption patterns and adapt heating, cooling, and lighting based on real-time occupancy, increasing thermal comfort for passengers while improving energy efficiency.

Additionally, CUSP leveraged digital twin capabilities for predictive modelling and scenario testing. Users could simulate strategies such as smart lighting controls and view projected energy savings before deployment. This improved decision-making and helped users select optimal energy-conservation measures.

Another essential practical contribution is the implementation of social media mining to extract information from citizens for energy management. Advanced natural language processing methods were developed to classify events, detect trends, and analyse sentiments from social data.

For example, by monitoring citizen satisfaction through sentiment analysis, building controls implemented within CUSP could identify and correct areas of thermal discomfort, showcasing how user-generated data can enable responsive and user-centric energy optimisation.

On a broader level, the thesis validates through prototypes and simulations the immense potential of combining semantic-driven AI, digital twins, and IoT for adaptable building energy management tuned precisely to usage and occupants.

Practical implementations underscore how semantics can provide buildings with intelligence to dynamically optimise comfort, efficiency, and sustainability. Evaluations demonstrate the ability of semantics to simplify complex systems and enhance usability.

In general, this research makes a notable contribution to the field of building energy management, demonstrating how the integration of semantics and AI within a BIM framework can result in improved energy efficiency and sustainability at the building level.

8.4.4 Using semantics at the level of a smart city

The fourth contribution is to the pioneering application of semantics in the context of smart cities, particularly by involving citizens as active sensors. The research highlights the transformative potential of incorporating semantically enhanced BIM data into smart city models, the key innovation being the active participation of citizens in the data collection process.

The thesis makes significant theoretical and practical contributions to the application of semantics at the city level to enhance user experience and decision-making.

On the theoretical front, the thesis proposes an innovative conceptual framework for leveraging citizens as "active sensors" within a smart city context. This approach posits that by interpreting and responding to data generated by citizens through their use of social media and other platforms, cities can become more dynamic, responsive, and "smart".

The thesis proposes a novel taxonomy to classify city events based on risk levels, facilitating real-time detection from social media data streams. This taxonomy was derived through an extensive review of the literature and an analysis of major smart city initiatives. The proposed event classification model enables automated risk assessment when events are detected.

The thesis presents an original methodology for the sentiment analysis of social media content to gauge citizen satisfaction with ongoing events and environmental factors. This approach provides a unique lens into the "pulse" of a smart city by tapping into its inhabitants' collective experiences and perspectives.

Multiple regression analysis establishes statistically significant correlations between event occurrences, sentiment trends, and external variables such as weather. These quantitative validations of the relationships between citizen satisfaction, urban events, and contextual factors represent vital theoretical contributions.

On the implementation side, the thesis details the development of an end-to-end pipeline for gathering, processing, classifying, analysing, and visualising data from social media and other sources. This stack encompasses the proposed taxonomy, NLP and machine learning techniques for classification, MRA for statistical analysis, and tools for geospatial representation.

The pipeline was deployed in a case study in Cardiff, UK, practically demonstrating the viability of semantically interpreting unstructured data from citizens to gain actionable insights. More than 1.6 million tweets were collected and classified, demonstrating the robustness and scalability of the implementation.

The interactive map developed allows city managers to visually explore detected events and sentiment hotspots, supporting data-driven decision making. The ability to simulate and evaluate different policy scenarios using historical data adds another layer of practical utility.

Another key conceptual contribution is the proposal to integrate BIM with semantics to create "context-informed digital twins" of buildings. Digital twins, as virtual replicas of physical buildings, can capture real-time data through IoT sensors. However, raw sensor data lack context or meaning. The thesis presents a novel approach to infusing semantics into the digital twin data using ontologies, allowing context to be derived from static and dynamic building artefacts.

For example, a temperature sensor reading is converted from a meaningless number to a contextualised insight that a particular zone or space may need cooling adjustments. This ability to interpret raw data in a contextual and human-centric way is a fundamental theoretical advance that bridges the gap between BIM, semantics, and user experience.

Furthermore, the thesis presents an innovative methodology to incorporate this semantically enriched digital twin into an energy management web application and to tailor the user interface and interactions based on contextual data. This approach aligns both the system functionality and the presentation of information with the specific user, his environment, tasks, and evolving needs.

Participatory research conducted with real-world stakeholders reveals key theoretical insights into how semantic digital twins can overcome complexities in built environments. By providing a more intuitive understanding of the interconnections between building components and energy systems, semantic digital twins theoretically improve usability, satisfaction, and decision making for users.

The integration of semantics also theoretically enables the personalisation of the energy management interface based on the user's context. For example, facility managers could receive higher-level metrics and controls, while tenants could receive targeted recommendations on reducing energy consumption. Such adaptability and context-awareness strengthen user engagement and satisfaction.

Furthermore, the use of iterative feedback loops in the participatory research process itself represents a key conceptual contribution. The cyclical process of participatory prototype refinement based on stakeholder perspectives aligns with the core philosophical paradigms underpinning the thesis. This approach resonates with the interpretivist stance of constructing meaning from subjective human experiences vs. objective truth.

In general, the novel integration of BIM, semantics, digital twins, and participatory research to enhance building energy management applications marks an important theoretical advancement. The human-centric digital representation of built environments has immense potential to revolutionise our understanding of and interaction with the spaces we occupy and manage.

The thesis also provides significant practical advancements in the use of semantics to improve user experience in smart city contexts.

A significant implementation contribution is the development of an end-to-end pipeline to gather, process, and analyse unstructured social media data from urban inhabitants. Practical modules include connectors for data ingestion, natural language processing models for classification, a semantic ontology, and a geospatial visualiser.

This pipeline was deployed for Cardiff, UK, practically demonstrating how semantics can be applied at the city scale by tapping into the "collective intelligence" of citizens as active data providers. The ability to handle real-world volumes of social data shows the scalability of the approach.

Specifically, the thesis implements an innovative event classification model using machine learning, enabling automated categorisation of incidents and happenings across the city based on the proposed risk taxonomy, providing immense practical utility for city managers to gain situational awareness.

Similarly, the sentiment analysis module provides practical value by identifying trends and patterns in emotional predilection about events, infrastructure, and services, as expressed through social media. City officials can take preemptive actions based on the early detection of negative sentiment trends.

The developed interactive visualiser practically demonstrates how all the semantically processed datasets can be visualised in a meaningful way for decision-makers. Users can explore trends and relationships spatially, such as through heatmaps, as well as through charts. This tool can help officials assess the impact of policies and plan interventions.

Participatory workshops conducted with city stakeholders provide helpful practical insight into how semantics can enhance the experience for users like urban planners and public works departments. Feedback helped refine prototypes to align with user needs, highlighting the applied value.

Using simulation features to create and assess different scenarios provides another layer of practical utility for smart city operations and planning. By learning from past trends, officials can estimate the outcomes of decisions before implementation.

Overall, the thesis provides substantive practical contributions demonstrating applied value - the pipeline implemented, modules developed, and participatory approach adopted converge to provide a real-world, semantics-based solution for enhancing the smart city user experience.

All in all, the amalgamation of conceptual models, methodological frameworks, and functional implementations concretely demonstrates the value of applying semantics at the city level. By semantically bridging the physical urban environment and its inhabitants, the thesis enables responsive and citizen-centric smart city solutions with both theoretical and practical advancements.

8.5 Limitations of the research

The study has made significant contributions in understanding and improving the user experience through semantics within smart city contexts, yet several limitations merit recognition and further exploration in future research.

First, the existing restricted literature on this subject has an inherent limitation. The research recognises that enhancing the user experience through semantics in smart city contexts is relatively uncharted territory, indicating that the study's findings and conclusions may hinge on a narrow corpus of existing literature.

Second, the scope of the research primarily focused on user experience enhancement is limited in the critical realm of security. This restricted focus might constrain the comprehensive understanding of the research, limiting its potential to provide a holistic view of the myriad factors shaping user experience in smart city environments.

A third limitation emanates from the rapidly evolving nature of web technologies. The rapid pace of advances in this field suggests that research, although relevant at present, can quickly become outdated as new technologies and trends emerge, posing challenges to its long-term applicability and relevance.

Fourth, the generalisability of the study findings presents another limitation. Research leans heavily on specific case studies or Web applications, which might restrict the broader applicability of the findings to other contexts or industries. Additionally, the external validity of the study may be compromised due to the limited sample size and the possible lack of diversity among participants involved in user interviews, surveys, and usability tests.

Technological constraints form a fifth limitation without detailed details, as alluded to in the study. These constraints could encompass availability or compatibility issues related to semantic web technologies, data sources, or computational resources, which could impact the execution and results of the research.

The sixth limitation pertains to contextual limitations. The research uses a digital twin to discuss the context derivation from static and dynamic artefacts. However, the complexities of accurately deriving context from these artefacts still need to be fully addressed. This limitation suggests potential challenges in accurately capturing and representing context, which could impact the effectiveness of user experience enhancement measures.

Finally, a fundamental limitation arises from the limited focus on security. Despite its primary emphasis on user experience, research needs to dive deeper into the complexities of security measures and privacy policies. This restricted focus may curtail the study's ability to comprehensively understand the security implications of enhancing user experience.

These limitations offer potential avenues for future research, necessitating further exploration to deepen the understanding of user experience in smart city contexts, keeping pace with rapidly evolving web technologies and expanding the study's scope to ensure it remains robust, relevant, and comprehensive.

Within the overarching narrative of the thesis, acknowledging the limitations of the bespoke UX model is as essential as recognising its substantial advancements in urban planning and management. Certain constraints come under scrutiny when scaling up semantics at the city level, with citizens acting as active sensors.

One primary challenge is Limited Direct User Engagement. In this phase, the participatory design process leaned heavily on the analysis of content posted by users on social networks, moving away from direct user interviews of previous stages. This change was a necessary adaptation to the wide and varied urban context. However, it introduced complications as the scale and diversity resulted in indirect user engagement. Consequently, this led to a dependency on inferred, as opposed to explicit, user feedback. This limitation could hinder the capture of nuanced needs and experiences within the user base, potentially overlooking important subjective user feedback.

The research presented in the thesis had a focused Scope of Research. It mainly focused on integrating semantics to enhance the user experience in specific areas, namely, BIM web applications, energy management platforms, and smart city data. However, additional studies are required to assess the broader applicability of these approaches in different systems and domains. The limited scope of the initial case studies can restrict the relevance of the findings to other contexts or industries, suggesting the need for a more expansive application of the research to validate its broader effectiveness.

Furthermore, evaluating the model's capacity to improve user experiences relied on relatively small Sample Sizes and Diversity. Users and experts involved in the studies may not represent broader user populations, affecting the external validity of the findings. More extensive empirical studies that include diverse participants are essential to establish the efficacy of semantic enhancements more conclusively. The geographical and cultural concentration of the research also underlines the importance of additional testing in various settings, reinforcing the need to affirm the global viability of the UX model.

Technological and Contextual Constraints also posed significant barriers. Issues related to the availability or compatibility of semantic web technologies, the diversity of data sources, or the limits of computational resources affect both the implementation and the results. The challenges related to deriving context accurately from static and dynamic artefacts within the digital twin model might remain unresolved; a gap raises questions about the capability to capture and model the complex nature of urban environments aptly.

Lastly, long-term robustness and data ethics are an area of concern. Although the methodologies designed were meant to be adaptive, their durability must be verified over time through extended research. Adopting the improved approaches depends on active stakeholder participation and the challenge of moving away from established legacy systems. Furthermore, increasing scrutiny around data ethics highlights the need for continuous review and refinement of data collection, retention, and sharing practices. By improving monitoring and audit methods, research could better align with evolving privacy standards and expectations.

In summary, despite the significant progress that the bespoke UX model showcases in enhancing user experiences through semantic integration in systems such as BIM, energy management, and smart cities, these limitations underscore the imperative for further investigation. Addressing these areas is critical to ensure that the model can effectively handle contemporary challenges. Additional research could improve the efficacy of the model, expand its relevance in various domains, and strengthen its resilience in a rapidly changing technological and urban landscape.

8.6 Future work

Although the thesis makes significant contributions to advance knowledge and practical implementations in the field of semantics to enhance user experience in smart cities and building energy management, it is not without limitations that provide fertile ground for further research. Specifically, restrictions exist regarding the scope of the literature, security considerations, rapidly evolving technologies, external validity, technical constraints, contextual modelling accuracy, and privacy implications. However, by addressing these limitations through extensive future work across a myriad of domains, the door remains open for considerable additional progress in employing semantics to unlock more intuitive, responsive, and tailored user experiences within our built environments and cities.

8.6.1 Expand the scope of the bespoke User Experience model

The bespoke UX model's journey towards creating a fully comprehensive and universally applicable approach remains underway, with substantial headway in integrating semantics into user experience within domains such as BIM, energy management, and smart city systems. Although the contributions made are significant, the identification of certain limitations throughout the research indicates opportunities for future enhancements, paving the way for a model that remains robust in the evolving tapestry of urban and technological environments and becomes more widely applicable across diverse paradigms.

Firstly, improving user engagement strategies emerges as a pivotal area for advancement. In light of the limited direct user engagement encountered thus far, future investigations must engineer innovative approaches to interact with users more intimately and effectively. Strategies could include developing interactive platforms to garner real-time user feedback or using virtual and augmented reality tools to create immersive user experiences. Further research efforts must also refine methods for collecting and analysing user feedback garnered indirectly, such as through social networks, to ensure that the gathered insights are not only accurately interpreted but translated into meaningful and actionable improvements.

Expanding the focus of research is another critical venture. The ambition is to elevate the relevance of the UX model from its current specialised applications to a broader arena of systems and domains. Venturing into disparate sectors such as healthcare, education, and transportation, which can significantly benefit from semantic enhancements in user experience, could amplify the model's relevance. Achieving this would necessitate forging collaborations with professionals from varied disciplines, pooling together wide-ranging perspectives and expert insights to enrich and diversify the model's capability.

Equally important is the need to support empirical studies' sample size and diversity. Future research should commission broader studies, drawing on a more extensive and varied pool of users and experts. This expansion is vital to augmenting the study's external validity, ensuring that conclusions drawn encompass and reflect diverse populations. The target population should include participants from distinct geographic and cultural backgrounds, fostering a globally nuanced understanding of the UX model's adaptability and application.

Technological and contextual constraints warrant meticulous attention in future work, as they entail a concerted effort to navigate and overcome issues about the availability, compatibility, and implementation of semantic web technologies, data sources, and computational resources. Investments in developing resilient and agile semantic technologies and developing cutting-edge data sourcing and computation methods are fundamental. Moreover, there is a crucial need to advance techniques for deriving accurate context from both static and dynamic artefacts within digital twins, which is crucial for capturing the essence of intricate and evolving urban ecosystems.

In conjunction with adherence to data ethics, the assurance of the model's longevity and robustness is indispensable. Longitudinal studies are necessary to affirm the enduring integrity of frameworks and to ensure that they can gracefully adapt and flourish amidst ongoing technological advancements. Constant monitoring and evaluation of the model performance in various settings will be integral. With the increasing discourse surrounding data ethics, future investigations must intensively review and improve the protocols involved in data collection, retention, and dissemination. Forging stringent oversight and audit mechanisms will be paramount to maintaining congruence with burgeoning privacy standards and ethical considerations.

Through the assiduous pursuit of these future research directions, the intention is not solely to fortify the model's efficacy and widen its scope of application, but also to guarantee its ongoing relevance and malleability within a landscape characterised by swift technological evolution and urban diversity. The path to this future is collaboration and transdisciplinarity, underscored by a commitment to innovation, inclusion, and unwavering ethical principles. It is on this path that semantics and user experience will harmoniously converge, ushering in a future where urban environments are smarter, more responsive, and more inclusively attuned to the needs and aspirations of their denizens.

8.6.2 Expanding the scope and scalability of the research

To enhance the breadth and impact of this research, more extensive efforts are needed to broaden its scope across disciplines, case studies, data sources, and demographics of participants. Greater scalability of the proposed solutions is also required for real-world implementation.

A crucial direction is to conduct more exhaustive reviews of the literature that cover related topics in semantics, user experience, BIM, smart cities, and adjacent fields. Current reviews have focused on specific sub-domains; a multi-disciplinary synthesis can provide a fuller theoretical basis and reveal additional gaps to be addressed. For instance, the potential for cross-pollination of concepts between construction informatics and human-computer interaction has not been thoroughly investigated. Comprehensive reviews would also benefit from analysing Semiotics, Cognitive Science, and Environmental Psychology literature to incorporate perspectives on human meaning-making, reasoning, and relation to the built environment.

Furthermore, evaluating the proposed methods and frameworks through more diverse and more extensive case studies can improve generalisability beyond the domains examined. Although this research has focused on the AEC industry as an application area, solutions should be tested for other data-intensive sectors such as healthcare, transportation, utilities and government. Comparison of results between industry verticals can highlight common principles in semantic UX design versus the domain-specific adaptations needed. Additionally, the implementation of case studies in different regions of the world can elucidate cultural and contextual parameters that influence the user experience.

Another crucial direction is to increase the sample size and expand the demographic diversity of the participants in user studies, interviews, surveys, focus groups, and usability tests. The human-centred nature of UX necessitates capturing a broad spectrum of user perspectives across attributes like age, profession, expertise, educational qualifications, socioeconomic status, and tech-savviness. Testing with larger and more heterogeneous user samples can improve generalisability and reveal subgroup differences in solution perception and adoption. For example, novices and casual users may prefer simpler interfaces than expert power users.

Furthermore, proposed solutions must be evaluated for applicability in diverse smart city environments, requiring pilot studies in multiple cities that differ in size, density, culture, climate, architectural styles, technological maturity, and governance models. Although certain characteristics may quickly transfer to new contexts, other location-specific adaptations would be required. For instance, visual interfaces for a North American city may require tweaks in colour schemes and iconography for an East Asian city. Such contextual testing can improve scalability and provide insights into customisation needs.

Finally, expanding the datasets used to develop and evaluate solutions is vital to scalability, ensuring that the models work with heterogeneous real-world data. Current methods rely on limited datasets; larger corpora, including multimedia, multilingual, and real-time streaming data, better reflect data diversity in the field. Additionally, increasing the use of sensor data, satellite imagery, geospatial data, and other alternate sources can enrich context information beyond the text-based sources currently dominating. More robust models trained on expanded datasets can effectively handle complex variables that influence UX in the built environment.

In summary, increasing the scope of the research through multifaceted literature analysis, extensive case studies, enlarged user samples, contextual evaluations, and enriched training data can elevate proposed solutions from isolated proofs of concept to large-scale, real-world impact. This expansion in breadth and depth can unlock the full potential of semantics to enhance user experiences within the built environment.

8.6.3 Expanding the scope and scalability of the research

To enhance the breadth and impact of this research, more extensive efforts are needed to broaden its scope across disciplines, case studies, data sources, and demographics of participants. Greater scalability of the proposed solutions is also required for real-world implementation.

A crucial direction is to conduct more exhaustive reviews of the literature that cover related topics in semantics, user experience, BIM, smart cities, and adjacent fields. Current reviews have focused on specific sub-domains; a multi-disciplinary synthesis can provide a fuller theoretical basis and reveal additional gaps to be addressed. For instance, the potential for cross-pollination of concepts between construction informatics and human-computer interaction has yet to be thoroughly investigated. Comprehensive reviews would also benefit from analysing semiotics, cognitive science, and environmental psychology literature to incorporate perspectives on human meaning-making, reasoning, and relation to the built environment.

Furthermore, evaluating the proposed methods and frameworks through more diverse and extensive case studies can improve generalisability beyond the domains examined. Although this research has focused on the AEC industry as an application area, solutions should be tested for other data-intensive sectors such as healthcare, transportation, utilities, and government. Comparing the results between industry verticals can highlight common principles in semantic UX design versus the domain-specific adaptations needed. Additionally, the implementation of case studies in different regions of the world can elucidate the cultural and contextual parameters that influence the user experience.

Another crucial direction is to increase the sample size and expand the demographic diversity of the participants in user studies, interviews, surveys, focus groups, and usability tests. The human-centred nature of UX necessitates capturing a broad spectrum of user perspectives across attributes like age, profession, expertise, educational qualifications, socioeconomic status, and tech-savviness. Testing with larger and more heterogeneous user samples can improve generalisability and reveal subgroup differences in solution perception and adoption. For example, novices and casual users may prefer simpler interfaces than expert power users.

Furthermore, proposed solutions must be evaluated for applicability in diverse smart city environments, requiring pilot studies in multiple cities that differ in size, density, culture, climate, architectural styles, technological maturity, and governance models. Although certain characteristics may quickly transfer to new contexts, other location-specific adaptations would be required. For instance, visual interfaces for a North American city may require tweaks in colour schemes and iconography for an East Asian city. Such contextual testing can improve scalability and provide insights into customisation needs.

Finally, expanding the datasets used to develop and evaluate solutions is vital to scalability, ensuring that the models work with heterogeneous real-world data. Current methods rely on limited datasets; larger corpora, including multimedia, multilingual, and real-time streaming data, better reflect data diversity in the field. Additionally, increasing the use of sensor data, satellite imagery, geospatial data, and other alternate sources can enrich context information beyond the text-based sources currently dominating. More robust models trained on expanded datasets can effectively handle complex variables that influence UX in the built environment.

In summary, increasing the scope of the research through multifaceted literature analysis, extensive case studies, enlarged user samples, contextual evaluations, and enriched training data can elevate proposed solutions from isolated proofs of concept to large-scale, real-world impact. This expansion in breadth and depth can unlock the full potential of semantics to enhance user experiences within the built environment.

8.6.4 Emerging concepts in the research

To maintain relevance and leverage leading-edge innovations, research should expand to cover emerging technological and conceptual paradigms shaping the built environment.

A key direction is to examine the potential of digital twins and BIM scanning to enhance user experiences for professionals and citizens interacting with urban data. Digital twins, virtual replicas of physical assets and processes, are gaining traction in architecture and construction. Integrating them with semantics could enable immersive and interactive experiences for collaborators to visualise and manipulate designs jointly. Scan-to-BIM facilitates the creation of as-built BIM models from 3D scan data, which is crucial for retrofits and heritage preservation. Incorporating semantics can enrich captured engineering data with historical and cultural meaning, engaging citizens.

Furthermore, integrating augmented, virtual and mixed reality (AR/VR/MR) with semantic models warrants investigation. These technologies are transforming spatial perception and interaction. Semantically annotated AR/VR/MR environments could contextualise digital overlays, improving comprehension of engineering data or heritage sites. User studies should evaluate how these technologies enhance understanding, collaboration, and decision making when combined with semantics.

Furthermore, investigating the interaction between semantics, predictive analytics, and optimisation algorithms can reveal new capabilities. Semantics can strengthen predictive analytics by providing contextual relationships between data variables. Predictions can, in turn, feed optimisation algorithms to identify ideal configurations for urban systems. This could enable intelligent, semantically contextualised optimisation to tackle challenges such as energy management.

Given its growing prominence in urban informatics, the scope of the research should also encompass emerging paradigms like ethics, sustainability, and social inclusion. Semantics could potentially encode ethical constraints, environmental impact data, and accessibility requirements into solutions. User studies are needed to evaluate the effectiveness of semantic interfaces in influencing positive community behaviour and values.

In summary, expanding the research focus to encompass leading-edge concepts like digital twins, immersive technologies, predictive optimisation, and ethical computing can position it at the frontier of innovation in the built environment. It can reveal novel high-impact semantics applications to enrich experiences and address complex urban challenges. Exploring these interdisciplinary connections will be crucial for the evolution of the field.

8.6.5 Assessing the impact and behaviour change enabled by the research

To fully gauge the value of the proposed solutions, a rigorous assessment of their impact in the real world is essential, including quantified usability improvements, the effectiveness in driving positive behaviour, and readiness for adoption.

A key direction is conducting controlled experiments to obtain quantifiable data on improvements in user experience metrics such as task completion times, workload, learnability, errors, and satisfaction. Comparing semantically enhanced interfaces with baselines will provide evidence of UX improvements attributed to semantics. The methodology should encompass diverse tasks, users, and environments to obtain nuanced insights. Qualitative feedback from users will complement quantitative data.

Additionally, the efficacy of semantically enriched interfaces in promoting positive behaviour change should be investigated through longitudinal studies and field trials. For example, its long-term effectiveness in promoting energy conservation, sustainability and community participation can be assessed by tracking related behaviours before and after deployment. Success indicators could encompass reduced energy consumption, increased recycling, or increased participation in community events. Control groups that do not receive semantic interventions can enable comparative analysis.

Furthermore, collaborative design workshops with domain experts can assess readiness for real-world adoption across factors such as relevance, usability, integration complexity, and deployment costs. Constructive feedback from architects, urban planners, policymakers, and technology specialists will highlight promising applications and remaining barriers, guiding further refinement. Comparative assessment can determine suitability across different solution components, features, and architectural configurations.

Lastly, large-scale longitudinal studies are imperative to understand long-term user engagement enabled by semantic UX design. Multi-year analysis can reveal patterns in sustained usage, evolving needs, and emergent practices as solutions integrate into everyday workflows. Periodic iterative refinement can align solutions with changing user expectations and technological landscapes, ensuring persistent relevance.

In summary, various evaluation strategies that encompass usability testing, impact analysis, expert reviews, and longitudinal follow-up can provide compelling multifaceted evidence of the effectiveness of enhanced semantic UX in facilitating positive transformation. Quantified assessments can strengthen the case for the adoption of semantic interfaces to improve productivity, conservation, inclusion, and overall quality of life in built environments.

The proposed multifaceted research directions encompass both depth and breadth, ranging from extensive literature analysis and large-scale user studies to nuanced technical refinements and evaluation of behaviour change. Collectively, these ambitious but integrated efforts can significantly advance the research frontier, unveiling new capabilities and applications that enrich experiences in the built environment. As solutions permeate real-world practice, sustained accumulation of evidence will solidify the integral role of semantics in shaping the next generation of intelligent, responsive, and human-centric cities. Building on the progress made so far, the next phase of this exciting journey promises profound innovation that elevates societies around the world.

With these expansive possibilities on the horizon, this research agenda lays the foundations for transformative change. As increasing urbanisation and technological disruption shape the built environment, the need for humanistic design has never been greater.

This research aims to place people at the heart of this change, using semantics to forge a harmonious convergence between technologies and experiences. The road ahead will undoubtedly present challenges, yet the promise of progress makes it worth travelling.

8.7 Rounding up

In conclusion, this thesis has explored the role of semantics in enhancing user experience for building information modelling (BIM) web applications. Through a rigorous three-stage research methodology that integrates literature reviews, gap analyses, prototype development, and continuous user feedback, the potential of semantics has been thoroughly investigated.

Stage I demonstrated how semantics can be leveraged to improve learning and training in the construction industry. The proposed approach for analysing textual and social media data revealed critical BIM roles and skills, emphasising the importance of socio-organisational aspects. Creating a dynamic semantic web platform ensured that these roles and skills remained up-to-date as BIM adoption evolved.

Stage II highlighted how semantics helps derive context from static and dynamic artefacts using digital twins. Integrating IoT sensors and semantic middleware, the CUSP platform architecture enabled context-aware data interpretation and intuitive user interfaces. Participatory evaluations validated the ability of CUSP to optimise energy performance when applied to complex urban infrastructures.

Stage III illustrated scaling up semantics at a city level with citizens as active sensors. Advanced natural language processing techniques were used to gain insight from social media data. Multiple regression analysis revealed interlinks between citizen satisfaction, events, and environmental factors. In general, the approach demonstrated the viability of harnessing collective intelligence to create inclusive, responsive smart cities.

Through its rigorous methodology and prototype implementations, this thesis has provided substantial evidence for the value of semantics in improving user experiences for BIM web applications. The research contributes significantly to the body of knowledge at the intersection of semantics, user experience, BIM, and smart cities. Practical recommendations have been offered to enhance decision-making, optimise design, and facilitate interactive simulations in the built environment.

However, certain limitations provide avenues for future work. The data sets used, although statistically significant, could be expanded for a more comprehensive analysis. Emerging semantic technologies, such as digital twins, warrant further research as they mature. Moreover, as fake news and cyber-threats proliferate, strategies to safeguard user experience must be explored. However, this thesis represents a robust framework and methodology to leverage semantics to create more intuitive, efficient, and meaningful interactions between users and BIM web applications.

In summary, this thesis has systematically investigated the potential of semantics to enhance user experience in building information modelling (BIM) web applications. Through an iterative three-stage approach, the research has demonstrated semantics' ability to enrich learning, optimise energy management, and engage citizens, ultimately transforming user interactions across the built environment.

Although limitations exist, they present opportunities for future work to build on these seminal explorations at the intersection of semantics, user experience, and BIM. As the pioneering journey of discovery continues, this thesis provides a solid foundation and methodology for researchers and practitioners who seek to create more intuitive, efficient, and meaningful digital experiences.

The transformative capability of semantics is evident; what remains is our collective imagination and effort to mould this capability into solutions that uplift individuals and communities. If this thesis inspires even incremental progress in that direction, it will have served its purpose. The possibilities at the confluence of technology, design, and human aspirations remain endless.

References

[1] Martina A. Clarke, Lisa L Neitzke, and Kathryn M Cooper. Usability of documentation software used by direct support providers: A proof-of-concept study. In *Human Factors in Aging and Special Needs*, AHFE, 2023. doi: 10.54941/ahfe1003672.

[2] F.H. Abanda, J.H.M. Tah, and R. Keivani. Trends in built environment semantic web applications: Where are we today? *Expert Syst. Appl*, 40(14):5563–5577, October 2013. ISSN 09574174. doi: 10.1016/j.eswa.2013.04.027.

[3] Gregory D Abowd and Bill N Schilit. Ubiquitous computing: The impact on future interaction paradigms and hci research. In *CHI'97 Extended Abstracts on Human Factors in Computing Systems*, pages 221–222, March 1997. ISBN 0897919262. doi: 10.1145/1120212.1120354.

[4] Ritu Agarwal and Jayesh Prasad. A conceptual and operational definition of personal innovativeness in the domain of information technology. *Information systems research*, 9(2):204–215, 1998. doi: 10.1287/isre.9.2.204. URL `https://doi.org/10.1287/isre.9.2.204`.

[5] A. Ahankoob, K. Manley, and B. Abbasnejad. The role of contractors' building information modelling (bim) experience in realising the potential values of bim. *International Journal of Construction Management*, 2019. doi: 10.1080/15623599.2019.1639126.

[6] Vian Ahmed, Algan Tezel, Zeeshan Aziz, and Magda Sibley. Big data in aec industry: A review of present status, opportunities, and future trends. *Facilities*, 35(13/14):725–745, October 2017. ISSN 0263-2772. doi: 10.1108/F-06-2016-0064.

[7] Yaser Hasan Al-Mamary and Alina Shamsuddin. Testing of the technology acceptance model in context of yemen. *Mediterranean Journal of Social Sciences*, 6(4):268, 2015. doi: 10.5901/mjss.2015.v6n4s1p268.

[8] Lauralee Alben. Quality of experience: defining the criteria for effective interaction design. *interactions*, 3(3):11–15, 1996. doi: 10.1145/235008.235010.

[9] Khalid Alemerien and Kenneth Magel. Slc: A visual cohesion metric to predict the usability of graphical user interfaces. In *Proceedings of the 30th Annual ACM Symposium on Applied Computing - SAC '15*, pages 1526–1533, April 2015. ISBN 9781450331968. doi: 10.1145/2695664.2695791.

[10] H Peter Alesso and Craig F Smith. *Thinking on the Web: Berners-Lee, Godel and Turing*. 2008.

[11] Khaled A Alshare, Hani I Mesak, Elizabeth E Grandon, and Masood A Badri. Examining the moderating role of national culture on an extended technology acceptance model. *Journal of*

Global Information Technology Management, 14(3):27–53, 2011. doi: 10.1080/1097198X.2011.10856542. URL https://doi.org/10.1080/1097198X.2011.10856542.

[12] Ole Andreas Alsos and Dag Svanæs. Designing for the secondary user experience. *Lecture Notes in Computer Science*, 6949:84–91, 2011. doi: 10.1007/978-3-642-23768-3_7.

[13] Mats Alvesson and K Skoldberg. Reflexive methodologies: Interpretation and research, 2000.

[14] Abdulrahman Alzahrani, Ali Alqazzaz, Nabil Almashfi, Huirong Fu, and Ye Zhu. Web application security tools analysis. *Studies in Media and Communication*, 5(2):118, November 2017. ISSN 2325-808X. doi: 10.11114/smc.v5i2.2663.

[15] Kwasi Amoako-Gyampah and Abdus F Salam. An extension of the technology acceptance model in an erp implementation environment. *Information & management*, 41(6):731–745, 2004. doi: j.im.2003.08.010. URL https://doi.org/10.1016/j.im.2003.08.010.

[16] Aleksandar Antonic, Martina Marjanovic, and Ivana Podnar vzarko. Modeling aggregate input load of interoperable smart city services. In *Proceedings of the 11th ACM International Conference on Distributed and Event-based Systems - DEBS '17*, pages 34–43, June 2017. ISBN 9781450350655. doi: 10.1145/3093742.3093928.

[17] Grigoris Antoniou and Frank van Harmelen. Web ontology language: Owl. *Handbook on ontologies*, pages 91–110, 2009. doi: 10.1007/978-3-540-92673-3_4.

[18] David Antons, Eduard Grunwald, Patrick Cichy, and Torsten Oliver Salge. The application of text mining methods in innovation research: current state, evolution patterns, and development priorities. *R&D Management*, 50(3):329–351, June 2020. ISSN 0033-6807. doi: 10.1111/radm.12408.

[19] Chid Apte, Fred Damerau, Sholom Weiss, et al. *Text mining with decision rules and decision trees*. 1998. doi: 10.1016/s0167-739x(97)00021-6.

[20] Gabor Aranyi and Paul Van Schaik. Modeling user experience with news websites. *Journal of the Association for Information Science and Technology*, 66(12):2471–2493, 2015. doi: 10.1002/asi.23348.

[21] Carmelo Ardito, Paolo Buono, Danilo Caivano, Maria Francesca Costabile, Rosa Lanzilotti, Anders Bruun, and Jan Stage. Usability evaluation: a survey of software development organizations. In *SEKE*, pages 282–287, 2011.

[22] Anonymous ASCE. Computational geometric approach for bim semantic enrichment to support automated underground garage compliance checking. *Journal of Construction Engineering and Management*, 148(1):05021013, January 2023. doi: 10.1061/(ASCE)CO.1943-7862.0002230. 2.

[23] Mohsen Attaran and Bilge Gokhan Celik. Digital twin: Benefits, use cases, challenges, and opportunities. *Decision Analytics Journal*, 6:100165, March 2023. ISSN 27726622. doi: 10.1016/j.dajour.2023.100165.

[24] Salman Azhar. Building information modeling (bim): Trends, benefits, risks, and challenges for the aec industry. *Leadership and Management in Engineering*, 11(3):241–252, July 2011. doi: 10.1061/(ASCE)LM.1943-5630.0000127.

[25] Salman Azhar, Abid Nadeem, Johnny YN Mok, and Brian HY Leung. Building information modeling (bim): A new paradigm for visual interactive modeling and simulation for construction projects. In *Proc., First International Conference on Construction in Developing Countries*, volume 1, pages 435–46, 2008.

[26] Franz Baader, Ian Horrocks, and Ulrike Sattler. Description logics as ontology languages for the semantic web. In Dieter Hutter, Werner Stephan, Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen, editors, *Mechanizing Mathematical Reasoning: Essays in Honor of Jorg H. Siekmann on the Occasion of His 60th Birthday*, volume 2605 of *Lecture notes in computer science*, pages 228–248. 2005. ISBN 978-3-540-25051-7. doi: 10.1007/978-3-540-32254-2_14.

[27] Kristina Baculakova et al. Selected aspects of smart city concepts: Position of bratislava. *Theoretical and Empirical Researches in Urban Management*, 15(3):68–80, 2020. doi: 10.1080/13604810802479126.

[28] Ricardo Baeza-Yates, Berthier Ribeiro-Neto, et al. *Modern information retrieval*, volume 463. 1999.

[29] M. Bano and D. Zowghi. User involvement in software development and system outcomes: a systematic literature review. *Requirements Engineering*, 18(4):289–327, 2013. doi: 10.1007/s00766-012-0172-5.

[30] Javier A. Bargas-Avila and Kasper Hornbaek. In the eye of the beholder: A conceptual model for user experience and semantic interactions. *International Journal of Human-Computer Studies*, 69(1-2):1–15, 2011. doi: 10.1016/j.ijhcs.2010.08.004.

[31] MB Barison and ET Santos. The competencies of bim specialists: a comparative analysis of the literature review and job ad descriptions. In *Computing in Civil Engineering (2011)*, pages 594–602. 2011. doi: 10.1061/41182(416)73.

[32] K. Battarbee and I. Koskinen. Co-experience: user experience as interaction. *CoDesign*, 1(1): 5–18, 2005. doi: 10.1080/15710880412331289917.

[33] Hartwig Baumgaertel, Hans Ehm, Sabrine Laaouane, Jan Gerhardt, and Anna Kasprzik. Collaboration in supply chains for development of cps enabled by semantic web technologies.

In *2018 Winter Simulation Conference (WSC)*, pages 3627–3638, December 2018. ISBN 978-1-5386-6572-5. doi: 10.1109/WSC.2018.8632455.

[34] TH Beach, Y Rezgui, and OF Rana. Cloudbim: Management of bim data in a cloud computing environment. *Cardiff School of Engineering, Cardiff University, Wales, UK*, 2011.

[35] Thomas Beach, Ioan Petri, Yacine Rezgui, and Omer Rana. Management of collaborative bim data by federating distributed bim models. *Journal of Computing in Civil Engineering*, 31(4): 04017009, 2017.

[36] Burcin Becerik-Gerber, Farrokh Jazizadeh, Nan Li, and Gulben Calis. Application areas and data requirements for bim-enabled facilities management. *Journal of construction engineering and management*, 138(3):431–442, 2012.

[37] Jakob Beetz, Jos van Leeuwen, and Bauke de Vries. Ifcowl: A case of transforming express schemas into ontologies. *Artificial intelligence for engineering design, analysis and manufacturing*, 23(1):89–101, 2009. doi: 10.1017/S0890060409000122.

[38] Emanuele Bellini, Pierfrancesco Bellini, Daniele Cenni, Paolo Nesi, Gianni Pantaleo, Irene Paoli, and Michela Paolucci. An ioe and big multimedia data approach for urban transport system resilience management in smart cities. *Sensors*, 21(2):435, 2021. doi: 10.3390/s21020435.

[39] D. Benyon. *Designing user experience*. 2019.

[40] Tim Berners-Lee. Long live the web. *Scientific American*, 303(6):80–85, December 2010. ISSN 0036-8733. doi: 10.1038/scientificamerican1210-80.

[41] Tim Berners-Lee, James Hendler, and Ora Lassila. The semantic web. *Scientific american*, 284 (5):34–43, May 2001. ISSN 0036-8733. doi: 10.1038/scientificamerican0501-34.

[42] Regina Bernhaupt. User experience evaluation in entertainment. *Human-Computer Interaction Series*, HCIS:3–7, 2009. doi: 10.1007/978-1-84882-963-3_1.

[43] Nigel Bevan, James Carter, and Susan Harker. Iso 9241-11 revised: What have we learnt about usability since 1998? In Masaaki Kurosu, editor, *International Conference on Human-Computer Interaction*, volume 9169 of *Lecture notes in computer science*, pages 143–151. Springer, 2015. ISBN 978-3-319-20900-5. doi: 10.1007/978-3-319-20901-2_13.

[44] Nigel Bevan, Jim Carter, Jonathan Earthy, Thomas Geis, and Susan Harker. New iso standards for usability, usability reports and usability measures. In Masaaki Kurosu, editor, *International conference on human-computer interaction*, volume 9731 of *Lecture notes in computer science*, pages 268–278. Springer, 2016. ISBN 978-3-319-39509-8. doi: 10.1007/978-3-319-39510-4_25.

[45] Simon Elias Bibri and John Krogstie. Smart sustainable cities of the future: An extensive interdisciplinary literature review. *Sustainable Cities and Society*, 31:183–212, 2018. doi: 10.1016/j.scs.2017.03.017.

[46] Muhammad Bilal, Lukumon O Oyedele, Junaid Qadir, Kamran Munir, Saheed O Ajayi, Olugbenga O Akinade, Hakeem A Owolabi, Hafiz A Alaka, and Maruf Pasha. Big data in the construction industry: A review of present status, opportunities, and future trends. *Advanced engineering informatics*, 30(3):500–521, August 2016. ISSN 14740346. doi: 10.1016/j.aei.2016.07.001.

[47] F Biljecki, H Ledoux, and J Stoter. A state-of-the-art review on the integration of building information modeling (bim) and geographic information system (gis). *ISPRS International Journal of Geo-Information*, 6(2):53, 2017. doi: 10.3390/ijgi6020053.

[48] F Biljecki, K Ali, J Wang, and H Ledoux. City digital twin potentials: A review and research agenda. *Sustainability*, 13(6):3386, 2021. doi: 10.3390/su13063386.

[49] Nimish Biloria. Infomatters, a multi-agent systems approach for generating performative architectural formations. *International Journal of Architectural Computing*, 13(2):134–150, 2015.

[50] Nimish Biloria. Urban (i)mpressions: Digitally driven urban metabolism framework for circular economy in built environments. In *Design Cybernetics*, pages 133–150, 2019.

[51] C. Bizer, Tom Heath, and Tim Berners-Lee. Linked data - the story so far. *Int. J. Semantic Web Inf. Syst*, 5(3):1–22, July 2009. ISSN 1552-6283. doi: 10.4018/jswis.2009081901.

[52] Harry E Blanchard. Update on recent hci and usability standards. *ACM SIGCHI Bulletin*, 30 (2):14–15, April 1998. ISSN 0736-6906. doi: 10.1145/279044.279047.

[53] A. Blandford, T. R. Green, D. Furniss, and S. Makri. The importance of identity and vision to user experience designers on agile projects. *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*, pages 4797–4802, 2009. doi: 10.1145/1520340.1520736.

[54] Barry W. Boehm. A spiral model of software development and enhancement. *Computer*, 21(5): 61–72, May 1988. ISSN 0018-9162. doi: 10.1109/2.59.

[55] Andrey Bogdanov and Takanori Isobe. *White-Box Cryptography Revisited: Space-Hard Ciphers*, page 1058–1069. October 2015. ISBN 9781450338325. doi: 10.1145/2810103.2813699.

[56] O Fals Borda, Peter Reason, and H Bradbury. Participatory (action) research in social theory: Origins and challenges. *Handbook of action research: Participative inquiry and practice*, pages 27–37, 2006.

[57] Pim Borst, Hans Akkermans, and Jan Top. Engineering ontologies. *International journal of human-computer studies*, 46(2-3):365–406, February 1997. ISSN 10715819. doi: 10.1006/ijhc.1996.0096.

[58] Conrad Boton, Sylvain Kubicki, and Gilles Halin. The challenge of level of development in 4d/bim simulation across aec project lifecyle. a case study. *Procedia Engineering*, 123:59–67, 2015. ISSN 18777058. doi: 10.1016/j.proeng.2015.10.058.

[59] Tim Bray, Jean Paoli, C. Michael Sperberg-McQueen, Eve Maler, and Francois Yergeau. Extensible markup language (xml) 1.0 (fifth edition). W3C Recommendation, 2008.

[60] Dan Brickley and R. V. Guha. Rdf schema 1.1. W3C Recommendation, 2014.

[61] J. Brier, C. M. Gray, and Y. Kou. In search of ux translators. In *Proceedings of the 2017 Conference on Designing Interactive Systems*, pages 957–962, 2017. doi: 10.1145/3064857.3079143.

[62] J Briskilal and CN Subalalitha. Classification of idiomatic sentences using awd-lstm. In *Expert Clouds and Applications*, pages 113–124. 2022. doi: 10.1007/978-981-16-2126-0_11.

[63] David Bryde, Marti Broquetas, and Jurgen Marc Volm. The project benefits of building information modelling (bim). *International journal of project management*, 31(7):971–980, October 2013. ISSN 02637863. doi: 10.1016/j.ijproman.2012.12.001.

[64] G Burrell. Morgan. g.(1979) sociological paradigms and organizational analysis. *London: I-Ieinemann*, 1979.

[65] Erik Cambria, Bjorn Schuller, Yunqing Xia, and Catherine Havasi. New avenues in opinion mining and sentiment analysis. *IEEE Intelligent systems*, 28(2):15–21, March 2013. ISSN 1541-1672. doi: 10.1109/MIS.2013.30.

[66] Brad Camburn, Varun Viswanathan, Julie Linsey, Derek Anderson, Daniel Jensen, Richard Crawford, and Kristin Wood. Design prototyping methods: State of the art in strategies, techniques, and guidelines. *Design Science*, 3:E13, 2017. doi: 10.1017/dsj.2017.10.

[67] R. Cameron. A sequential mixed model research design: Design, analytical and display issues. *Int. J. Mult. Res. Approaches*, 3(2):140–152, August 2009. ISSN 1834-0806. doi: 10.5172/mra.3.2.140.

[68] Edgar M Cano-Cruz and Francisco Lopez-Orozco. Design and development of a low-cost and meteorological system: Meteoblue. *Advanced Science Letters*, 21(1):83–87, 2015. doi: 10.1166/asl.2015.5760.

[69] Veronica Castaneda, Luciana Ballejos, Ma Laura Caliusco, and Ma Rosa Galli. The use of ontologies in requirements engineering. *Global journal of researches in engineering*, 10(6):2–8, 2010.

[70] Manuel Castells. *The Internet Galaxy: Reflections on the Internet, Business, and Society.* 2001.

[71] CMT Castillo-Calzadilla, K Zabala, Eneko Arrizabalaga, P Hernandez, L Mabe, JR Lopez, JM Casado, MN Santos, J Guardo, B Molinete, et al. The opportunity for smart city projects at municipal scale: Implementing a positive energy district in zorrozaurre. *Ekonomiaz*, 2021.

[72] Pew Research Center. *Social Media Fact Sheet.* 2018.

[73] Tomo Cerovsek. A review and outlook for a 'building information model'(bim): A multi-standpoint framework for technological development. *Advanced engineering informatics*, 25(2): 224–244, April 2011. ISSN 14740346. doi: 10.1016/j.aei.2010.06.003.

[74] Junyi Chai, James N.K. Liu, and Eric W.T. Ngai. Application of decision-making techniques in supplier selection: A systematic review of literature. *Expert Systems with Applications*, 40 (10):3872–3885, 2013. doi: 10.1016/10.1016/j.eswa.2012.12.040.

[75] Moulay Larbi Chalal, Medjdoub Benachir, Michael White, and Raid Shrahily. Energy planning and forecasting approaches for supporting physical improvement strategies in the building sector: A review. *Renewable and sustainable energy reviews*, 64:761–776, October 2016. ISSN 13640321. doi: 10.1016/j.rser.2016.06.040.

[76] Hock Chuan Chan and Hock-Hai Teo. Evaluating the boundary conditions of the technology acceptance model. *ACM Transactions on Computer-Human Interaction*, 14(2):9, 2007. doi: 10.1145/1275511.1275515.

[77] Hung-Ming Chen, Kai-Chuan Chang, and Tsung-Hsi Lin. A cloud-based system framework for performing online viewing, storage, and analysis on big data of massive bims. *Automation in Construction*, 71:34–48, November 2016. ISSN 09265805. doi: 10.1016/j.autcon.2016.03.002.

[78] Zeyu Chen, Lei Yang, Jue Wu, and Fujun Yang. A multi-threaded virtual dom difference method based on web worker. In Shijie Jia and Huajun Dong, editors, *Sixth International Conference on Computer Information Science and Application Technology (CISAT 2023)*, October 2023. doi: 10.1117/12.3003797.

[79] Paul Chynoweth, Sharon Christensen, Judith McNamara, and Kathryn O'Shea. Legal and contracting issues in electronic project administration in the construction industry. *Structural Survey*, 2007.

[80] Cities. Sustainable cities and society. *Sustainable cities and society*, 62:102352, 2020.

[81] Jurgen Cito, Gerald Schermann, John Erik Wittern, Philipp Leitner, Sali Zumberi, and Harald c Gall. An empirical analysis of the docker container ecosystem on github. In *2017 IEEE/ACM 14th International Conference on Mining Software Repositories (MSR)*, pages 323–333. IEEE, May 2017. ISBN 978-1-5386-1544-7. doi: 10.1109/MSR.2017.67.

[82] Tim Clerckx, Jan Van den Bergh, and Karin Coninx. Modeling multi-level context influence on the user interface. In *Fourth Annual IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOMW'06)*, pages 5–pp. IEEE, 2006.

[83] Andrew W Coburn, Gary Bowman, Simon J Ruffle, Roxane Foulser-Piggott, Daniel Ralph, and Michelle Tuveson. A taxonomy of threats for complex risk management. *Cambridge Risk Framework series*, 2014.

[84] Christopher Colvin and Aileen Goh. Validation of the technology acceptance model for police. *Journal of Criminal Justice*, 33(1):89–95, 2005. doi: 10.1016/j.jcrimjus.2004.10.009.

[85] S. Consoli. *Producing Linked Data for Smart Cities: the case of Catania*, volume 7. March 2014. doi: 10.1016/j.bdr.2016.10.001.

[86] Vincenzo Corrado, Ilaria Ballarini, Leandro Madrazo, and German Nemirovskij. Data structuring for the ontological modelling of urban energy systems: The experience of the semanco project. *Sustainable Cities and Society*, 14:223–235, February 2015. ISSN 22106707. doi: 10.1016/j.scs.2014.09.006.

[87] Cardiff City Council. Cardiff smart city roadmap, February 2020.

[88] Edoardo Croci and Tania Molteni. Business models for smart city solutions: An overview of main archetypes. *International Journal of Urban Planning and Smart Cities (IJUPSC)*, 2(2): 94–109, 2021. doi: 10.4018/IJUPSC.2021070106.

[89] D. Alan Cruse. Meaning in language: An introduction to semantics and pragmatics. *Oxford Textbooks in Linguistics*, 2011. doi: 10.1093/acprof:oso/9780199559466.001.0001. [Online; accessed 29. Jun. 2023].

[90] Edward Curry, James O'Donnell, Edward Corry, Souleiman Hasan, Marcus Keane, and Sean O'Riain. Linking building data in the cloud: Integrating cross-domain building data using linked data. *Advanced Engineering Informatics*, 27(2):206–219, 2013. doi: 10.1016/j.aei.2012.10.003.

[91] Tiago Silva Da Silva, Angela Martin, Frank Maurer, and Milene Silveira. User-centered design and agile methods: a systematic review. In *2011 Agile Conference*, pages 77–86. IEEE, 2011.

[92] Andrew Dainty, David Moore, and Michael Murray. *Communication in construction: Theory and practice*. January 2007. ISBN 9780203358641. doi: 10.4324/9780203358641.

[93] Wilfried Daniels, Danny Hughes, Mahmoud Ammar, Bruno Crispo, Nelson Matthys, and Wouter Joosen. *SM V - the security microvisor: A virtualisation-based security middleware for the internet of things*, page 36–42. December 2017. ISBN 9781450352000. doi: 10.1145/3154448.3154454.

[94] Moumita Das, Jack C.P. Cheng, and Kincho H. Law. An ontology-based web service framework for construction supply chain collaboration and management. *Engineering, Construction and Architectural Management*, 22(5):551–572, September 2015. ISSN 0969-9988. doi: 10.1108/ECAM-07-2014-0089.

[95] N.K. Dash. Research methods resource - selection of the research paradigm and methodology. *IGNOU*, 2005.

[96] Amos David and Nadine Ndjock. Big data, knowledge organization and decision making: opportunities and limits. In Fernanda Ribeiro and Maria Elisa Cerveira, editors, *Challenges and Opportunities for Knowledge Organization in the Digital Age*, pages 95–102. Ergon-Verlag, 2018. ISBN 9783956504211. doi: 10.5771/9783956504211-95.

[97] A Davies, T Brady, M Hobday, A Prencipe, A Tsang, and R DeFillippi. Building information modellingin construction: insights from collaboration and change management perspectives. *International Journal of Construction Management*, 17(2):136–147, 2017. doi: 10.1080/09537287.2017.1407005.

[98] J. Davies, R. Studer, and P. Warren. *Semantic Web technologies: trends and research in ontology-based systems*. 2006.

[99] Matteo Del Giudice and Anna Osello. Bim for cultural heritage. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 5(2):225–229, July 2013. ISSN 2194-9034. doi: 10.5194/isprsarchives-XL-5-W2-225-2013.

[100] Deloitte. Global mobile consumer trends, 2017.

[101] D.E.S.I. Connectivity - broadband market developments in the eu, 2019.

[102] Branko Dimeski, Memet Memeti, and Mimoza Bogdanoska-Jovanoska. International cooperation of the city of skopje: Projects for accelerating smart city developments. *Smart Cities and Regional Development (SCRD) Journal*, 3(1):89–101, 2019.

[103] L.Y. Ding, B.T. Zhong, S. Wu, and H.B. Luo. Construction risk knowledge management in bim using ontology and semantic web technology. *Safety science*, 87:202–213, August 2016. ISSN 09257535. doi: 10.1016/j.ssci.2016.04.008.

[104] F. Dinis, Joao Pocas Martins, and B. Rangel. Bim and semantic enrichment methods and applications: A review of recent developments. *Archives of Computational Methods in Engineering*, May 2021.

[105] ISO DIS. 9241-210: 2010. ergonomics of human system interaction-part 210: Human-centred design for interactive systems. *International Standardization Organization (ISO). Switzerland*, 2009.

[106] Soussan Djamasbi, Marisa Siegel, and Tom Tullis. Generation y, web design, and eye tracking. *International Journal of Human Computer Studies*, 68:307–323, 2010. ISSN 10715819. doi: 10.1016/j.ijhcs.2009.12.006.

[107] M. Dobreva, A. O'Dwyer, and P. Feliciati. Blackboard vs. moodle: Comparing user experience of learning management systems. In *2007 37th Annual Frontiers In Education Conference - Global Engineering:Knowledge Sharing And Networking*, pages T2G–7. IEEE, 2007. doi: 10.1109/FIE.2007.4417910.

[108] Anil R Doshi, Sharat Raghavan, Rebecca Weiss, and Eric Petitt. The impact of the supply of fake news on consumer behavior during the 2016 us election. *SSRN Electronic Journal*, 2018. ISSN 1556-5068. doi: 10.2139/ssrn.3093397.

[109] Yehezkel Dror. Muddling through-" science" or inertia? *Public administration review*, pages 153–157, 1964. doi: 10.2307/973640.

[110] Jing Du, Qi Zhu, Yangming Shi, Qi Wang, Yingzi Lin, and Daniel Zhao. Cognition digital twins for personalized information systems of smart cities: Proof of concept. *Journal of Management in Engineering*, 36(2):04019052, 2020. doi: 10.1061/me.2020.36.issue-2;ctype:string:Journal.

[111] Robert Eadie, Mike Browne, Henry Odeyinka, Clare McKeown, and Sean McNiff. Bim implementation throughout the uk construction project lifecycle: An analysis. *Automation in construction*, 36:145–151, December 2013. ISSN 09265805. doi: 10.1016/j.autcon.2013.09.001.

[112] Chuck Eastman, Paul Teicholz, Rafael Sacks, and Kathleen Liston. *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors*. 2011.

[113] European Construction Technology Platform (ECTP). Strategic research agenda for the european construction sector, 2005.

[114] Haya El-Ghalayini, Mohammed Odeh, and Richard McClatchey. Developing ontology-driven conceptual data models. In *Proceedings of the 1st International Conference on Intelligent Semantic Web-Services and Applications*, pages 1–6, June 2010. ISBN 9781450304757. doi: 10.1145/1874590.1874594.

[115] Rasha F Elgazzar and Rania El-Gazzar. Smart cities, sustainable cities, or both. In *A Critical Review and Synthesis of Success and Failure Factors. SMARTGREENS 2017 Conference*, pages 250–257, April 2017. ISBN 978-989-758-241-7. doi: 10.5220/0006307302500257.

[116] Lien Faelens, Kristof Hoorelbeke, Bart Soenens, Kyle Van Gaeveren, Lieven De Marez, Rudi De Raedt, and Ernst HW Koster. Social media use and well-being: A prospective experience-sampling study. *Computers in Human Behavior*, 114:106510, 2021. doi: 10.1016/j.chb.2020.106510.

[117] Laurent Fasnacht. mmappickle: Python 3 module to store memory-mapped numpy array in pickle format. *Journal of Open Source Software*, 3(26):651, June 2018. ISSN 2475-9066. doi: 10.21105/joss.00651.

[118] Ling Feng, Christopher Pineda Monterola, and Yanqing Hu. The simplified self-consistent probabilities method for percolation and its application to interdependent networks. *New journal of physics*, 17(6):063025, June 2015. ISSN 1367-2630. doi: 10.1088/1367-2630/17/6/063025.

[119] Eva Samaniego Fernandez, Maria Sol Velasco Sacristan, and Pedro Antonio Fuertes Olivera. Translations we live by: The impact of metaphor translation on target systems. In *8th International Cognitive Linguistics Conference*, volume 23, 2003.

[120] Abdul Karim Feroz, Hangjung Zo, and Ananth Chiravuri. Digital transformation and environmental sustainability: A review and research agenda. *Sustainability*, 13(3):1530, February 2021. ISSN 2071-1050. doi: 10.3390/su13031530.

[121] J. R. Firth. A synopsis of linguistic theory, 1930-1955. In *Studies in Linguistic Analysis*, 1957.

[122] Romano Fistola, Marco Raimondo, and Rosa Anna La Rocca. The smart city and mobility: The functional polarization of urban flow. In *2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, pages 532–537, June 2017. ISBN 978-1-5090-6484-7. doi: 10.1109/MTITS.2017.8005730.

[123] Jodi Forlizzi and Shannon Ford. The building blocks of experience: An early framework for interaction designers. In Daniel Boyarski and Wendy A. Kellogg, editors, *Proceedings of the conference on Designing interactive systems processes, practices, methods, and techniques - DIS '00*, pages 419–423, August 2000. ISBN 1581132190. doi: 10.1145/347642.347800.

[124] George Forman, Evan Kirshenbaum, and Jaap Suermondt. Pragmatic text mining: minimizing human effort to quantify many issues in call logs. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 852–861. ACM, 2006. doi: 10.1145/1150402.1150520.

[125] Aidan Fuller, Zhong Fan, Charles Day, and Chris Barlow. Digital twin: Enabling technologies, challenges and open research. *IEEE access*, 8:108952–108971, 2020. ISSN 2169-3536. doi: 10.1109/ACCESS.2020.2998358.

[126] Lu Gao and Neil Eldin. Employers' expectations: A probabilistic text mining model. *Procedia Engineering*, 85:175–182, 2014. doi: 10.1016/j.proeng.2014.10.542.

[127] Anna Hviid Garrett, Janet Nielsen, Thomas A Sick Nielsen, and Sonja Haustein. Free-floating carsharing in copenhagen: A study on user experience in a cycling city. *Danish Journal of Transportation Research–Dansk Tidsskrift for Transportforskning*, 3:14–34, 2021.

[128] Tsvetanka Georgieva-Trifonova, Kaloyan Zdravkov, and Donika Valcheva. Application of semantic technologies in bibliographic databases. *Ecos de Linguagem*, 38(1):113–137, December 2019. ISSN 0264-0473. doi: 10.1108/EL-03-2019-0081.

[129] Pierpaolo Girardi and Andrea Temporelli. Smartainability: a methodology for assessing the sustainability of the smart city. *Energy Procedia*, 111:810–816, March 2017. ISSN 18766102. doi: 10.1016/j.egypro.2017.03.243.

[130] M. Govindarajan and M. Govindarajan. Approaches and applications for sentiment analysis: A literature review. *IGI Global*, January 0001.

[131] Rebecca Grant. Archivists managing research data? a survey of irish organisations. *International Journal of Digital Curation*, 15(1):9, December 2020. ISSN 1746-8256. doi: 10.2218/ijdc.v15i1.693.

[132] C. M. Gray, A. L. Toombs, and T. Grossman. It's more of a mindset than a method": Ux practitioners' conception of design methods. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 4044–4055, 2016. doi: 10.1145/2858036.2858410.

[133] C. M. Gray, Y. Kou, B. Battles, J. Hoggatt, and A. L. Toombs. The dark (patterns) side of ux design. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–14. ACM, 2018. doi: 10.1145/3173574.3174108.

[134] Nancy B Grimm, Stanley H Faeth, Nancy E Golubiewski, Charles L Redman, Jianguo Wu, Xuemei Bai, and John M Briggs. Global change and the ecology of cities. *Science*, 319(5864): 756–760, February 2008. doi: 10.1126/science.1150195.

[135] G. Groger and L. Plumer. Citygml – interoperable semantic 3d city models. *ISPRS J. Photogramm. Remote Sens*, 71:12–33, 07 2012. ISSN 09242716. doi: 10.1016/j.isprsjprs.2012.04.004.

[136] Thomas R Gruber. Toward principles for the design of ontologies used for knowledge sharing? *International journal of human-computer studies*, 43(5-6):907–928, November 1995. ISSN 10715819. doi: 10.1006/ijhc.1995.1081.

[137] T.R. Gruber. A translation approach to portable ontology specifications. *Knowl. Acquis*, 5(2): 199–220, June 1993. ISSN 10428143. doi: 10.1006/knac.1993.1008.

[138] Yuan Xiang Gu, Harold Johnson, Clifford Liem, Andrew Wajs, and Michael J. Wiener. *White-box cryptography: Practical protection on hostile hosts*, page 1–8. December 2016. ISBN 9781450348416. doi: 10.1145/3015135.3015139.

[139] Hui Guan, Hongji Yang, and Hossam Hakeem. *Reverse engineering web applications for security mechanism enhancement*, page 492–497. July 2014. ISBN 978-1-4799-3578-9. doi: 10.1109/COMPSACW.2014.82.

[140] Nicola Guarino. Formal ontology and information systems. In *Proceedings of FOIS*, volume 98, pages 81–97, 1998.

[141] Mehmet Guccyetmez and Husham Sakeen Farhan. Enhancing smart grids with a new iot and cloud-based smart meter to predict the energy consumption with time series. *Alexandria Engineering Journal*, 79:44–55, September 2023. ISSN 11100168. doi: 10.1016/j.aej.2023.07.071.

[142] R. Guha and R. McCool. Semantic search. In *Proceedings of the 12th International Conference on World Wide Web*, pages 700–709, May 2003. ISBN 1581136803. doi: 10.1145/775152.775250.

[143] R.V. Guha, Dan Brickley, and Steve Macbeth. Schema.org: Evolution of structured data on the web. *Communications of the ACM*, 59(2):44–51, January 2016. ISSN 00010782. doi: 10.1145/2844544.

[144] Amanda Gutmanis. A collaborative governance process: The city of vancouver's 2014-2017 healthy city for all leadership table. In *Simon Fraser University Library*, 2022.

[145] Amelie Gyrard, Christian Bonnet, Karima Boudaoud, and Martin Serrano. Assisting iot projects and developers in designing interoperable semantic web of things applications. In *2015 IEEE International Conference on Data Science and Data Intensive Systems*, pages 659–666. IEEE, December 2015. ISBN 978-1-5090-0214-6. doi: 10.1109/DSDIS.2015.60.

[146] Beate Hampe and Joseph E Grady. *From perception to meaning: Image schemas in cognitive linguistics*, volume 29. 2005.

[147] Jan Hartmann, Alistair Sutcliffe, and Antonella De Angeli. Towards a theory of user judgment of aesthetics and user interface quality. *Acm Transactions on Computer-Human Interaction*, 2008. doi: 10.1145/1460355.1460357.

[148] M. Hassenzahl and N. Tractinsky. Hedonic/pragmatic model of user experience. *Towards a UX Manifesto*, pages 10–14, 2006. doi: 10.5555/1137351.1137357.

[149] Marc Hassenzahl. The thing and i: understanding the relationship between user and product. In *Funology: from usability to enjoyment*, pages 1–12. 2003. doi: 10.1007/1-4020-2967-5_4.

[150] Marc Hassenzahl. User experience (ux): towards an experiential perspective on product quality. *Proceedings of the 20th International Conference of the Association Francophone d'Interaction Homme-Machine*, pages 11–15, 2008. doi: 10.1145/1512714.1512717.

[151] Marc Hassenzahl. Experience design: Technology for all the right reasons. *Synthesis Lectures on Human-Centered Informatics*, 2010. doi: 10.2200/s00261ed1v01y201003hci008.

[152] Wei He and Lu Yang. Using wikis in team collaboration: A media capability perspective. *Information & Management*, 53(7):846–856, November 2016. ISSN 03787206. doi: 10.1016/j.im.2016.06.009.

[153] Jeffrey Heer and Ben Shneiderman. Interactive dynamics for visual analysis: A taxonomy of tools that support the fluent and flexible use of visualizations. *Queue*, 10(2):30–55, February 2012. ISSN 1542-7749. doi: 10.1145/2133416.2146416.

[154] Christiane M. Herr and Thomas Fischer. Bim adoption across the chinese aec industries: An extended bim adoption model. *J. Comput. Des. Eng.*, 6(2):173–178, April 2019. ISSN 2288-5048. doi: 10.1016/j.jcde.2018.06.001.

[155] Lorenz M Hilty, Peter Arnfalk, Lorenz Erdmann, Jim Goodman, Matthias Lehmann, and Patrick A Wager. The relevance of information and communication technologies for environmental sustainability - a prospective simulation study. *Environmental Modelling & Software*, 21(11):1618–1629, November 2006. ISSN 13648152. doi: 10.1016/j.envsoft.2006.05.007.

[156] Sarah El Himer, Mariyam Ouaissa, Mariya Ouaissa, Moez Krichen, Mohannad Alswailim, and Mutiq Almutiq. Energy consumption monitoring system based on iot for residential rooftops. *Computation*, 11(4):78, April 2023. ISSN 2079-3197. doi: 10.3390/computation11040078.

[157] J-L Hippolyte, Yacine Rezgui, Haijiang Li, Bejay Jayan, and Shaun Howell. Ontology-driven development of web services to support district energy applications. *Automation in Construction*, 86:210–225, February 2018. ISSN 09265805. doi: 10.1016/j.autcon.2017.10.004.

[158] Pascal Hitzler. A review of the semantic web field. *Communications of the ACM*, 64(2):76–83, January 2021. ISSN 0001-0782. doi: 10.1145/3397512.

[159] Andrei Hodorog, Ali Hussain S Alhamami, and Ioan Petri. Social media mining for bim skills and roles for energy efficiency. In *2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, pages 1–10. IEEE, 2019. doi: 10.1109/ICE.2019.8792571.

[160] Richard Holden and Ben-Tzion Karsh. The technology acceptance model: Its past and its future in health care. *Journal of Biomedical Informatics*, 43(1):159–172, 2010. doi: 10.1016/j.jbi.2009.07.002.

[161] Jeremy Howard and Sylvain Gugger. Fastai: a layered api for deep learning. *Information*, 11 (2):108, February 2020. ISSN 2078-2489. doi: 10.3390/info11020108.

[162] Jeremy Howard and Sebastian Ruder. Universal language model fine-tuning for text classification. *arXiv preprint arXiv:1801.06146*, pages 328–339, 2018. doi: 10.18653/v1/P18-1031.

[163] I-Ching Hsu. Multilayer context cloud framework for mobile web 2.0: a proposed infrastructure. *International Journal of Communication Systems*, 26(5):610–625, May 2013. ISSN 1074-5351. doi: 10.1002/dac.1365.

[164] Xiping Hu, Xitong Li, Edith C-H Ngai, Victor CM Leung, and Philippe Kruchten. Multidimensional context-aware social network architecture for mobile crowdsensing. *IEEE Communications Magazine*, 52(6):78–87, June 2014. ISSN 0163-6804. doi: 10.1109/MCOM.2014.6829948.

[165] Ying Hui Huang, Guan Yu Li, and Xin Long Li. A development model of semantic web application. *Advanced materials research*, 765-767:1275–1280, September 2013. ISSN 1662-8985. doi: 10.4028/www.scientific.net/AMR.765-767.1275.

[166] Trevor Hunter, Peter Worthy, Stephen Viller, and Ben Matthews. Designing for citizen-centric smart cities: Supporting people's needs of a community garden. In *Proceedings of the 30th Australian Conference on Computer-Human Interaction - OzCHI '18*, pages 162–166, December 2018. ISBN 9781450361880. doi: 10.1145/3292147.3292217.

[167] Ke Huo, Vinayak Karthik, and Ramani. Window-shaping: 3d design ideation in mixed reality. In *Proceedings of the 2016 Symposium on Spatial User Interaction - SUI '16*, pages 189–189, October 2016. ISBN 9781450340687. doi: 10.1145/2983310.2989189.

[168] Jorn Hurtienne, Kerstin Klockner, Sarah Diefenbach, Claudia Nass, and Andreas Maier. Designing with image schemas: resolving the tension between innovation, inclusion and intuitive use. *Interacting with Computers*, 27(3):235–255, May 2015. ISSN 0953-5438. doi: 10.1093/iwc/iwu049.

[169] ICAM. Icam architecture part ii-volume iv - function modeling manual (idef0), 1982.

[170] Valeriu Manuel Ionescu. The analysis of the performance of rabbitmq and activemq. In *2015 14th RoEduNet International Conference-Networking in Education and Research (RoEduNet NER)*, pages 132–137. IEEE, September 2015. ISBN 978-1-4673-8179-6. doi: 10.1109/RoEduNet.2015.7311982.

[171] Hassan Ismkhan and Mohammad Izadi. Proposing a dimensionality reduction technique with an inequality for unsupervised learning from high-dimensional big data. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 53(6):3880–3889, June 2023. ISSN 2168-2216. doi: 10.1109/TSMC.2023.3234227.

[172] ISO13407 ISO. 13407: Human-centred design processes for interactive systems. *Geneva: ISO*, 1999.

[173] Krzysztof Janowicz, Martin Raubal, and Werner Kuhn. The semantics of similarity in geographic information retrieval. *Journal of Spatial Information Science*, pages 29–57, May 2011. ISSN 1948-660X. doi: 10.5311/JOSIS.2011.2.3.

[174] X Jiang, G Liu, and c Deng. An application framework of digital twin and its case study. *Journal of Ambient Intelligence and Humanized Computing*, 10(11):4065–4077, 2018. doi: 10.1007/s12652-018-0911-3.

[175] Ying Jing, Jiaxing Cui, Ding Ma, and Yiyun Chen. Towards sustainable urban development: Use of geographic big data for spatial planning. *Frontiers in Environmental Science*, 11:552, 2023.

[176] Mark Johnson. *The body in the mind: The bodily basis of meaning, imagination, and reason.* 2013.

[177] Igor Jurisica, John Mylopoulos, and Eric Yu. Ontologies for knowledge management: an information systems perspective. *Knowledge and Information systems*, 6(4):380–401, July 2004. ISSN 0219-1377. doi: 10.1007/s10115-003-0135-4.

[178] Maxim Kalinin, Vasiliy Krundyshev, and Peter Zegzhda. Cybersecurity risk assessment in smart city infrastructures. *Machines*, 9(4):78, 2021. doi: 10.3390/machines9040078.

[179] Evangelos Kanoulas, Mihai Lupu, Paul Clough, Mark Sanderson, Mark Hall, Allan Hanbury, and Elaine Toms. *Information Access Evaluation–Multilinguality, Multimodality, and Interaction: 5th International Conference of the CLEF Initiative, CLEF 2014, Sheffield, UK, September 15-18, 2014, Proceedings*, volume 8685. 2014.

[180] Rita Karam, Susan G. Straus, Albert Byers, Courtney A. Kase, and Matthew Cefalu. The role of online communities of practice in promoting sociotechnical capital among science teachers. *Educational Technology Research and Development*, 66(2):1–31, September 2017. ISSN 1042-1629. doi: 10.1007/s11423-017-9541-2.

[181] Ebrahim P. Karan and Javier Irizarry. Developing a spatial data framework for facility management supply chains. In *Construction Research Congress 2014*, pages 2355–2364, May 2014. ISBN 9780784413517. doi: 10.1061/9780784413517.239.

[182] Ebrahim P. Karan, Javier Irizarry, and John Haymaker. Bim and gis integration and interoperability based on semantic web technology. *Journal of Computing in Civil Engineering*, 30(3): 04015043, May 2016. ISSN 0887-3801. doi: 10.1061/(ASCE)CP.1943-5487.0000519.

[183] Evangelos Karapanos, John Zimmerman, Jodi Forlizzi, and Jean-Bernard Martens. User experience over time: an initial framework. In *CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 729–738. April 2009. ISBN 978-1-60558246-7. doi: 10.1145/1518701.1518814.

[184] Mohamad Kassem, Abd Raoff, Nur Liyana, and Djamel Ouahrani. Identifying and analyzing bim specialist roles using a competency-based approach. In *Creative Construction Conference 2018, CCC 2018, 30 June - 3 July 2018, Ljubljana, Slovenia*, 2018. doi: 10.3311/ccc2018-135.

[185] P. Ketikidis, T. Dimitrovski, L. Lazuras, and P. Bath. Acceptance of health information technology in health professionals: An application of the revised technology acceptance model. *Health Informatics Journal*, 18(2):124–134, 2012. doi: 10.1177/1460458211435425.

[186] Muhammad Asif Hossain Khan, Danushka Bollegala, Guangwen Liu, and Kaoru Sezaki. Multi-tweet summarization of real-time events. In *2013 International Conference on Social Computing*, pages 128–133. IEEE, September 2013. ISBN 978-0-7695-5137-1. doi: 10.1109/SocialCom.2013.26.

[187] Rohit Khare and cCelik Tantek. Microformats: A pragmatic path to the semantic web. In *Proceedings of the 15th International Conference on World Wide Web*, pages 865–866, May 2006. ISBN 1595933239. doi: 10.1145/1135777.1135917.

[188] Qutaiba Khasawneh, Jennifer Dworak, Ping Gui, Benjamin Williams, Alan C. Elliott, and Anand Muthaiah. Real-time monitoring of test fallout data to quickly identify tester and yield issues in a multi-site environment. In *2018 IEEE 36th VLSI Test Symposium (VTS)*, pages 1–6, April 2018. ISBN 978-1-5386-3774-6. doi: 10.1109/VTS.2018.8368661.

[189] Dongwook Kim and Sungbum Kim. Role and challenge of technology toward a smart sustainable city: topic modeling, classification, and time series analysis using information and communication technology patent data. *Sustainable Cities and Society*, 82:103888, July 2022. ISSN 22106707. doi: 10.1016/j.scs.2022.103888.

[190] Kylie Goodell King. Measuring teamwork and team performance in collaborative work environments. *Evidence-based HRM: a Global Forum for Empirical Scholarship*, 5(2):196–205, August 2017. ISSN 2049-3983. doi: 10.1108/EBHRM-11-2016-0028.

[191] R. Kitchin. The real-time city? big data and smart urbanism. *GeoJournal*, 79(1):1–14, 2014. doi: 10.1007/s10708-013-9516-8.

[192] Rob Kitchin and Martin Dodge. The (in) security of smart cities: Vulnerabilities, risks, mitigation, and prevention. *Journal of Urban Technology*, 26(2):47–65, 2019. doi: 10.1080/10630732.2017.1408002.

[193] Inge M Klopping and Earl McKinney. Extending the technology acceptance model and the task-technology fit model to consumer e-commerce. *Information Technology, Learning & Performance Journal*, 22(1):35–48, 2004.

[194] B. P. Knijnenburg and M. C. Willemsen. Recommender systems: from algorithms to user experience. *User Modeling and User-Adapted Interaction*, 22(4-5):331–382, 2012. doi: 10.1007/s11257-011-9112-x.

[195] Mario J Kofler, Christian Reinisch, and Wolfgang Kastner. A semantic representation of energy-related information in future smart homes. *Energy and Buildings*, 47:169–179, April 2012. ISSN 03787788. doi: 10.1016/j.enbuild.2011.11.044.

[196] Nicos Komninos, Charalampos Bratsas, Christina Kakderi, and Panagiotis Tsarchopoulos. Smart city ontologies: Improving the effectiveness of smart city applications. *Journal of Smart Cities (Transferred)*, 1(1), November 2016. ISSN 23826401. doi: 10.18063/JSC.2015.01.001.

[197] Oliver Kramer. Scikit-learn. In *Machine learning for evolution strategies*, volume 20 of *Studies in big data*, pages 45–53. 2016. ISBN 978-3-319-33381-6. doi: 10.1007/978-3-319-33383-0_5.

[198] Werner Kritzinger, Matthias Karner, Georg Traar, Jan Henjes, and Wilfried Sihn. Digital twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*, 51 (11):1016–1022, 2018. ISSN 24058963. doi: 10.1016/j.ifacol.2018.08.474.

[199] Andreas Kruger and Thomas H Kolbe. Building analysis for urban energy planning using key indicators on virtual 3d city models-the energy atlas of berlin. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 39:145–150, July 2012. ISSN 2194-9034. doi: 10.5194/isprsarchives-XXXIX-B2-145-2012.

[200] S. Kujala. The role of user involvement in requirements quality and project success. *ICIS 2005 Proceedings*, pages 75–84, September 2005. doi: 10.1109/RE.2005.72.

[201] M. Kuniavsky. *Smart things: ubiquitous computing user experience design*. 2010.

[202] Carine Lallemand, Guillaume Gronier, and Vincent Koenig. User experience: A concept without consensus? exploring practitioners' perspectives through an international survey. *Computers in Human Behavior*, 43(43):35–48, February 2015. ISSN 07475632. doi: 10.1016/j.chb.2014.10.048.

[203] Manika Lamba and Margam Madhusudhan. Application of sentiment analysis in libraries to provide temporal information service: a case study on various facets of productivity. *Social Network Analysis and Mining*, 8(1):63, 2018. doi: 10.1007/s13278-018-0541-y.

[204] Craig Larman and Victor R Basili. Iterative and incremental developments. a brief history. *Computer*, 36(6):47–56, June 2003. ISSN 0018-9162. doi: 10.1109/MC.2003.1204375.

[205] E. L. C. Law and P. Van Schaik. Scaling the usability metric for user experience (sumi). *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction-Volume 1*, pages 215–221, 2008. doi: 10.14236/ewic/HCI2008.27.

[206] E. L. C. Law and P. Van Schaik. Scalability of the user experience. *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design*, pages 548–551, 2014. doi: 10.1145/2686612.2686713.

[207] E. L. C. Law, V. Roto, M. Hassenzahl, A. P. Vermeeren, and J. Kort. Understanding, scoping and defining user experience: a survey approach. In *Proceedings of the SIGCHI conference on human factors in computing systems*, volume 29, pages 719–728. ACM, 2009. doi: 10.1145/1518701.1518813.

[208] Irving Lazar, Richard Darlington, Harry Murray, Jacqueline Royce, Ann Snipper, and Craig T Ramey. Lasting effects of early education: A report from the consortium for longitudinal studies. *Monographs of the society for research in child development*, 47(2/3):i–151, 1982. ISSN 0037976X. doi: 10.2307/1165938.

[209] Albert Lederer, D.J. Maupin, and Y. Zhuang. The technology acceptance model and the world wide web. *Decision Support Systems*, 29(3):269–282, 2000. doi: 10.1016/s0167-9236(00)00076-2.

[210] A. Lee, L. Thi, and M. Lin. Affective technology acceptance model: Extending technology acceptance model with positive and negative affect. In *IntechOpen*, 2017. doi: 10.5772/intechopen.70351.

[211] Pengfei Li, Shengqin Zheng, Hongyun Si, and Ke Xu. Critical challenges for bim adoption in small and medium-sized enterprises: evidence from china. *Advances in Civil Engineering*, 2019: 1–14, February 2019. ISSN 1687-8086. doi: 10.1155/2019/9482350.

[212] Xia Li, Patrick SW Fong, Shengli Dai, and Yingchun Li. Towards sustainable smart cities: An empirical comparative assessment and development pattern optimization in china. *Journal of Cleaner Production*, 215:730–743, 2019. doi: 10.1016/j.jclepro.2019.01.046.

[213] Shin Liao, Jon-Chao Hong, Ming-Hui Wen, Yi-Chen Pan, et al. Applying technology acceptance model (tam) to explore users' behavioral intention to adopt a performance assessment system for e-book production. *EURASIA Journal of Mathematics, Science and Technology Education*, 14(10):em1601, 2018. doi: 10.29333/ejmste/93575. URL `https://doi.org/10.29333/ejmste/93575`.

[214] Yongxin Liao, Mario Lezoche, Herve Panetto, and Nacer Boudjlida. Semantic annotations for semantic interoperability in a product lifecycle management context. *International Journal of Production Research*, 54(18):5534–5553, September 2016. ISSN 0020-7543. doi: 10.1080/00207543.2016.1165875.

[215] Thaiana Lima, Rodrigo Pereira dos Santos, Jonice Oliveira, and Claudia Werner. The importance of socio-technical resources for software ecosystems management. *Journal of Innovation in Digital Ecosystems*, 3(2):98–113, December 2016. ISSN 23526645. doi: 10.1016/j.jides.2016.10.006.

[216] Shukun Liu, Weijia Jia, and Xianmin Pan. Fault-tolerant feedback virtual machine deployment based on user-personalized requirements. *Frontiers of Computer Science*, 12(4):682–693, August 2018. ISSN 2095-2228. doi: 10.1007/s11704-017-6422-0.

[217] Jorge Lopez-Castroman, Bilel Moulahi, Jerome Aze, Sandra Bringay, Julie Deninotti, Sebastien Guillaume, and Enrique Baca-Garcia. Mining social networks to improve suicide prevention: A scoping review. *Journal of neuroscience research*, 2019. doi: 10.1002/jnr.24404.

[218] De-cheng Lou and Tian-fang Yao. Semantic polarity analysis and opinion mining on chinese review sentences [j]. *Journal of Computer Applications*, 11:30–45, 2006.

[219] Lian Leng Low, Adlina Maulod, and Kheng Hock Lee. Evaluating a novel integrated community of care (icoc) for patients from an urbanised low-income community in singapore using the participatory action research (par) methodology: a study protocol. *BMJ Open*, 7(10):e017839, October 2017. ISSN 2044-6055. doi: 10.1136/bmjopen-2017-017839.

[220] Boratto Ludovico and Vargiu Eloisa. Data-driven user behavioral modeling: from real-world behavior to knowledge, algorithms, and systems. *Journal of Intelligent Information Systems*, 54(1):1–4, 2020.

[221] Nicholas H. Lurie and Charlotte H. Mason. Visual representation: Implications for decision making. *Journal of Marketing*, 71(1):160–177, January 2007. ISSN 1547-7185. doi: 10.1509/jmkg.71.1.160.

[222] John Lyons. *Linguistic Semantics: An Introduction.* November 1995. ISBN 9780511810213. doi: 10.1017/CBO9780511810213.

[223] Miltiadis D Lytras, Wadee Al-Halabi, Jacky Xi Zhang, Mehedi Masud, and Ramzi A Haraty. Enabling technologies and business infrastructures for next generation social media: Big data, cloud computing, internet of things and virtual reality. *J. Univers. Comput. Sci.*, 21(11): 1379–1384, 2015.

[224] Miltiadis D Lytras, Anna Visvizi, and Jari Jussila. Social media mining for smart cities and smart villages research, 2020.

[225] Elena Malakhatka, Liridona Sopjani, and Per Lundqvist. Co-creating service concepts for the built environment based on the end-user daily activities analysis: Kth live-in-lab explorative case study. *Sustainability*, 13(4):1942, February 2021. ISSN 2071-1050. doi: 10.3390/su13041942.

[226] Timothy Malche, Priti Maheshwary, and Rakesh Kumar. Environmental monitoring system for smart city based on secure internet of things (iot) architecture. *Wireless Personal Communications*, 107(4):2143–2172, 2019. doi: 10.1007/s11277-019-06376-0.

[227] M. Malik, R. Chaturvedi, and A. Kaushik. Machine learning in big data handling: Challenges and prospects. In *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, pages 2426–2431. IEEE, April 2022. doi: 10.1109/ICACITE53722.2022.9823793.

[228] Lynne C Manzo and Nathan Brightbill. Toward a participatory ethics. In *Participatory action research approaches and methods*, pages 59–66. 2007.

[229] Elwin Marg. Descartes' error: emotion, reason, and the human brain. *Optometry and Vision Science*, 72(11):847–848, November 1995. ISSN 1040-5488. doi: 10.1097/00006324-199511000-00013.

[230] Noortje Marres and Sarah de Rijcke. From indicators to indicating interdisciplinarity: A participatory mapping methodology for research communities in-the-making. *Quantitative Science Studies*, 1(3):1041–1055, September 2020. ISSN 2641-3337. doi: 10.1162/qss_a_00062.

[231] David Martin et al. Bringing semantics to web services: The owl-s approach. In Jorge Cardoso, Amit Sheth, David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Dough Tygar, Moshe Y. Vardi, and Gerhard Weikum, editors, *Proceedings of the First International Workshop on Semantic Web Services and Web Process Composition*, volume 3387 of *Lecture notes in computer science*, pages 26–42, 2004. ISBN 978-3-540-24328-1. doi: 10.1007/978-3-540-30581-1_4.

[232] Jean Everson Martina. *The Ever Changing Threat Model: A Social-Technical Perspective*, page 1–1. June 2013. ISBN 978-0-7695-5065-7. doi: 10.1109/STAST.2013.16.

[233] Pedro Martins, Daniel Albuquerque, Cristina Wanzeller, Filipe Caldeira, Paulo Tome, and Filipe Sa. Cityaction a smart-city platform architecture. In *Advances in information and communication: proceedings of the 2019 future of information and communication conference (FICC), volume 1*, volume 69 of *Lecture notes in networks and systems*, pages 217–236. 2020. ISBN 978-3-030-12387-1. doi: 10.1007/978-3-030-12388-8_16.

[234] Sebastian Marwecki, Maximilian Brehm, Lukas Wagner, Lung-Pan Cheng, Florian'Floyd' Mueller, and Patrick Baudisch. Virtualspace-overloading physical space with multiple virtual reality users. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–10, April 2018. ISBN 9781450356206. doi: 10.1145/3173574.3173815.

[235] Trisna Maulida, Agnis Nur Afa Zumaroh, Hasri Akbar Awal Rozaq, Alif Yahya Syafa'at, Rona Sepri Ananda, Imam Tahyudin, Bambang Pilu Hartato, and Berlilana. Visualization of front-end data logger internet of things technology using vue.js framework. In *2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, pages 693–698, 2022. doi: 10.1109/ICITISEE57756.2022.10057919.

[236] Bright Gameli Mawudor, Myong-Hee Kim, and Man-Gon Park. *Continuous monitoring methods as a mechanism for detection and mitigation of growing threats in banking security system*, page 1–5. December 2015. ISBN 978-1-5090-1669-3. doi: 10.1109/IDM.2015.7516317.

[237] Kris McGlinn, Baris Yuce, Hendro Wicaksono, Shaun Howell, and Yacine Rezgui. Usability evaluation of a web-based tool for supporting holistic building energy management. *Automation*

in Construction, 84:154–165, December 2017. ISSN 0926-5805. doi: 10.1016/j.autcon.2017.08.033.

[238] Deborah L. McGuinness and Frank Van Harmelen. Owl web ontology language overview. W3C Recommendation, 2004.

[239] Jude Chua Soo Meng. Donald schon, herbert simon and the sciences of the artificial. *Design Studies*, 30(1):60–68, 2009.

[240] Ines Mergel. Distributed democracy: Seeclickfix. com for crowdsourced issue reporting. *Com for Crowdsourced Issue Reporting (January 27, 2012)*, 2012. ISSN 1556-5068. doi: 10.2139/ssrn.1992968.

[241] Stephen Merity, Nitish Shirish Keskar, and Richard Socher. Regularizing and optimizing lstm language models. *arXiv preprint arXiv:1708.02182*, 2017.

[242] Farid Meziane and Yacine Rezgui. A document management methodology based on similarity contents. *Information Sciences*, 158:15–36, 2004. doi: 10.1016/j.ins.2003.08.009.

[243] Reijo Miettinen and Sami Paavola. Beyond the bim utopia: Approaches to the development and implementation of building information modeling. *Automation in construction*, 43:84–91, July 2014. ISSN 09265805. doi: 10.1016/j.autcon.2014.03.009.

[244] Jenna Mikus, Deanna Grant-Smith, and Janice Rieger. *Cultural Probes as a Carefully Curated Research Design Approach to Elicit Older Adult Lived Experience*, page 182–207. December 2021. doi: 10.4018/978-1-7998-8479-8.ch009.

[245] Aliaksei Miniukovich, Simone Sulpizio, and Antonella De Angeli. Visual complexity of graphical user interfaces. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces - AVI '18*, pages 1–9, May 2018. ISBN 9781450356169. doi: 10.1145/3206505.3206549.

[246] Alexander G Mirnig, Alexander Meschtscherjakov, Daniela Wurhofer, Thomas Meneweger, and Manfred Tscheligi. A formal analysis of the iso 9241-210 definition of user experience. In *Proceedings of the 33rd annual ACM conference extended abstracts on human factors in computing systems*, pages 437–450, April 2015. ISBN 9781450331463. doi: 10.1145/2702613.2732511.

[247] Rohan Mishra, Pradyumn Prakhar Sinha, Ramit Sawhney, Debanjan Mahata, Puneet Mathur, and Rajiv Ratn Shah. Snap-batnet: Cascading author profiling and social network graphs for suicide ideation detection on social media. In *Proceedings of the 2019 conference of the North American Chapter of the association for computational linguistics: student research workshop*, pages 147–156, 2019.

[248] Indra Mohan, Amar Kumar Das, Ramakrishnan Prakash, and Sachin Kumar. Prediction modeling using artificial neural network (ann) for the performance and emission characteristics of catalytic co-pyrolytic fuel blended with diesel in a ci engine. *Environmental Science and Pollution Research*, pages 1–14, October 2023. ISSN 1614-7499. doi: 10.1007/s11356-023-30209-0.

[249] Jae Moon and Young-Gul Kim. Extending the tam for a world-wide-web context. *Information & Management*, 38(4):217–230, 2001. doi: 10.1016/s0378-7206(00)00061-6.

[250] Naomi Morishita-Steffen, Remi Alberola, Baptiste Mougeot, Etienne Vignali, Camilla Wikstrom, Uwe Montag, Emmanuel Gastaud, Brigitte Lutz, Gerhard Hartmann, Franz Xaver Pfaffenbichler, et al. Smarter together: Progressing smart data platforms in lyon, munich, and vienna. *Energies*, 14(4):1075, February 2021. ISSN 1996-1073. doi: 10.3390/en14041075.

[251] Peter Morville. *Ambient findability: What we find changes who we become.* 2005.

[252] Myriam Munezero, Calkin Suero Montero, Erkki Sutinen, and John Pajunen. Are they different? affect, feeling, emotion, sentiment, and opinion detection in text. *IEEE transactions on affective computing*, 5(2):101–111, 2014.

[253] Mike Murray. *Rethinking construction: the egan report (1998).* January 2003. ISBN 9780470758526. doi: 10.1002/9780470758526.ch13.

[254] Sam Musa. Smart cities - a roadmap for development. *Journal of Telecommunications System & Management*, 5(3):1–3, 2016.

[255] Ganesh Nalluru, Rahul Pandey, and Hemant Purohit. Relevancy classification of multimodal social media streams for emergency services. In *2019 IEEE International Conference on Smart Computing (SMARTCOMP)*, pages 121–125. IEEE, June 2019. ISBN 978-1-7281-1689-1. doi: 10.1109/SMARTCOMP.2019.00040.

[256] N. O. Nawari, T. Chichugova, S. Mansoor, and L. Delfin. Bim in structural design education. In *Computing in Civil and Building Engineering (2014)*, pages 2143–2150, June 2014. ISBN 9780784413616. doi: 10.1061/9780784413616.266.

[257] Timothy Neate, Aikaterini Bourazeri, Abi Roper, Simone Stumpf, and Stephanie Wilson. Co-created personas: Engaging and empowering users with diverse needs within the design process. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*, pages 1–12, May 2019. ISBN 9781450359702. doi: 10.1145/3290605.3300880.

[258] Paolo Nesi, Claudio Badii, Pierfrancesco Bellini, Daniele Cenni, Giacomo Martelli, and Michela Paolucci. Km4city smart city api: an integrated support for mobility services. In *2016 IEEE International Conference on Smart Computing (SMARTCOMP)*, pages 1–8. IEEE, May 2016. ISBN 978-1-5090-0898-8. doi: 10.1109/SMARTCOMP.2016.7501702.

[259] Eric WT Ngai and Philtip Tin Yun Lee. A review of the literature on applications of text mining in policy making. In *PACIS*, page 343, 2016.

[260] J. Nielsen. Usability inspection methods. In *Conference companion on Human factors in computing systems*, pages 413–414. ACM, 1994. doi: 10.1145/259963.260531.

[261] Jakob Nielsen. *The Usability Engineering Lifecycle*. 1993. doi: 10.1016/b978-0-08-052029-2.50007-3.

[262] Lene Nielsen. *Personas - User Focused Design*. 2013. doi: 10.1007/978-1-4471-4084-9.

[263] Raja Shahmir Nizam, Cheng Zhang, and Lu Tian. A bim based tool for assessing embodied energy for buildings. *Energy and Buildings*, 170:1–14, July 2018. ISSN 0378-7788. doi: 10.1016/j.enbuild.2018.03.067.

[264] Kunihiro Noda, Takashi Kobayashi, Tatsuya Toda, and Noritoshi Atsumi. *Identifying core objects for trace summarization using reference relations and access analysis*, page 13–22. July 2017. ISBN 978-1-5386-0367-3. doi: 10.1109/COMPSAC.2017.142.

[265] D. A. Norman. *The Design of Everyday Things: Revised and Expanded Edition*. 2013.

[266] Don Norman. Emotion & design: attractive things work better. *Interactions*, 9(4):36–42, July 2002. ISSN 1072-5520. doi: 10.1145/543434.543435.

[267] Don Norman, Jim Miller, and Austin Henderson. What you see, some of what's in the future, and how we go about doing it: Hi at apple computer. In I. Katz, R. Mack, and L. Marks, editors, *Conference companion on Human factors in computing systems*, page 155. ACM, May 1995. ISBN 0897917553. doi: 10.1145/223355.223477.

[268] Donald A. Norman. *Emotional design: Why we love (or hate) everyday things*. 2004.

[269] Donald A Norman and Stephen W Draper. *User centered system design: New perspectives on human-computer interaction*. January 1986. ISBN 9781482229639. doi: 10.1201/b15703.

[270] Daniel Oberle, Stephan Grimm, and Steffen Staab. An ontology for software. In *Handbook on ontologies*, pages 383–402. 2009.

[271] Timothy O Olawumi, Abdullahi B Saka, Daniel WM Chan, and Nimesha S Jayasena. Scientometric review and analysis: A case example of smart buildings and smart cities. In Emmanuel Manu and Julius Akotia, editors, *Secondary Research Methods in the Built Environment*, pages 147–162. March 2021. ISBN 9781003000532. doi: 10.1201/9781003000532-11.

[272] Rosamaria Olivadese, Beril Alpagut, Beatriz Pineda Revilla, Jeroen Brouwer, Vasiliki Georgiadou, Alexander Woestenburg, and Mark van Wees. Towards energy citizenship for a just and inclusive transition: Lessons learned on collaborative approach of positive energy districts from

the eu horizon2020 smart cities and communities projects. In *Multidisciplinary Digital Publishing Institute Proceedings*, volume 65-1, page 20, January 2021. doi: 10.3390/proceedings2020065020.

[273] Charles O'Reilly and Andrew JM Binns. The three stages of disruptive innovation: Idea generation, incubation, and scaling. *California Management Review*, 61(3):49–71, 2019. ISSN 2162-8564. doi: 10.1177/0008125619841878.

[274] W.J. Orlikowski and J.J. Baroudi. Studying information technology in organizations: Research approaches and assumptions. *Inf. Syst. Res*, 2(1):1–28, March 1991. ISSN 1047-7047. doi: 10.1287/isre.2.1.1.

[275] R. Orpwood, C. Gibbs, T. Adlam, R. Faulkner, and D. Meegahawatte. The design of smart homes for people with dementia-user-interface aspects. *Universal Access in the Information Society*, 4:156–164, 2005. doi: 10.1007/s10209-005-0120-7.

[276] D. Ostrowski, N. Rychtyckyj, P. MacNeille, and M. Kim. Integration of big data using semantic web technologies, February 2016.

[277] Timo Partala and Veikko Surakka. The effects of affective interventions in human–computer interaction. *Interacting With Computers*, 16(2):295–309, 2004. doi: 10.1016/j.intcom.2003.12.001.

[278] Pieter Pauwels and Anna Roxin. Simplebim: From full ifcowl graphs to simplified building graphs. In *eWork and eBusiness in Architecture, Engineering and Construction*, pages 11–18. 2017.

[279] Pieter Pauwels and Walter Terkaj. Express to owl for construction industry: Towards a recommendable and usable ifcowl ontology. *Automation in Construction*, 63:100–133, 03 2016. ISSN 09265805. doi: 10.1016/j.autcon.2015.12.003.

[280] Pieter Pauwels, Sijie Zhang, and Yong-Cheol Lee. Semantic web technologies in aec industry: A literature overview. *Automation in Construction*, 73:145–165, January 2017. ISSN 09265805. doi: 10.1016/j.autcon.2016.10.003.

[281] G.C. Peng, J.M.B. Nunes, and F. Annansingh. Investigating information systems with mixed-methods research. In *Proceedings of the IADIS International Workshop on Information Systems Research Trends, Approaches and Methodologies*. Sheffield, 2011.

[282] Veridiana Rotondaro Pereira, Melanie E. Kreye, and Marly Monteiro de Carvalho. Customer-pulled and provider-pushed pathways for product-service system. *Journal of Manufacturing Technology Management*, 30(4):729–747, January 2019. ISSN 1741-038X. doi: 10.1108/JMTM-07-2018-0209.

[283] Stephen G. Perz, Marliz Arteaga, Andrea Baudoin Farah, I. Foster Brown, Elsa Renee Huaman Mendoza, Yara Araujo Pereira de Paula, Leonor Mercedes Perales Yabar, Alan dos Santos Pimentel, Sabina C. Ribeiro, Guillermo Rioja-Ballivian, Martha Cecilia Rosero Pena, L. Cecilia Sanjinez L., and N. Galia Selaya G. Participatory action research for conservation and development: Experiences from the amazon. *Sustainability*, 14(1):233, December 2021. ISSN 2071-1050. doi: 10.3390/su14010233.

[284] Ioan Petri, Tom Beach, Yacine Rezgui, Ian E Wilson, and Haijiang Li. Engaging construction stakeholders with sustainability through a knowledge harvesting platform. *Computers in Industry*, 65(3):449–469, 2014. doi: 10.1016/j.compind.2014.01.008.

[285] Ioan Petri, Haijiang Li, Yacine Rezgui, Yang Chunfeng, Baris Yuce, and Bejay Jayan. A modular optimisation model for reducing energy consumption in large scale building facilities. *Renewable and Sustainable Energy Reviews*, 38:990–1002, October 2014. ISSN 13640321. doi: 10.1016/j.rser.2014.07.044.

[286] Ioan Petri, Omer F Rana, Tom Beach, and Yacine Rezgui. Performance analysis of multi-institutional data sharing in the clouds4coordination system. *Computers & Electrical Engineering*, 58:227–240, 2017. doi: 10.1016/j.compeleceng.2017.02.015.

[287] Ioan Petri, Ateyah Alzahrani, Jonathan Reynolds, and Yacine Rezgui. Federating smart cluster energy grids for peer-to-peer energy sharing and trading. *IEEE Access*, 8:102419–102435, 2020. doi: 10.1109/ACCESS.2020.2998747.

[288] H. Petrie and N. Bevan. *The evaluation of accessibility, usability, and user experience*, volume 1. 2009. doi: 10.1201/9781420064995-c20.

[289] J. Ploennigs, H. Diblowski, A. Roder, K. Kabitzsch, and B. Hensel. Deliverable d4.1: Ict platform for holistic energy efficiency simulation and lifecycle management of public use facilities, 2011.

[290] Jim Plume and John Mitchell. Collaborative design using a shared ifc building model–learning from experience. *Automation in Construction*, 16(1):28–36, January 2007. ISSN 09265805. doi: 10.1016/j.autcon.2005.10.003.

[291] J. Powell. Building information modelling (bim), 2013.

[292] P. Pu, L. Chen, and R. Hu. Explaining the user experience of recommender systems. *User Modeling and User-Adapted Interaction*, 22(4-5):441–484, 2012. doi: 10.1007/s11257-011-9118-4.

[293] P. Pu, L. Chen, and R. Hu. A user-centric evaluation framework for recommender systems. *User Modeling and User-Adapted Interaction*, 22(4-5):349–382, 2012. doi: 10.1007/s11257-011-9117-5.

[294] Francesco Pucillo and Gaetano Cascini. A framework for user experience, needs and affordances. *Design Studies*, 35(2):160–179, March 2014. ISSN 0142694X. doi: 10.1016/j.destud.2013.10.001.

[295] Hilary Putnam. Meaning and reference. *The journal of philosophy*, 70(19):699–711, November 1974. ISSN 0022362X. doi: 10.2307/2025079.

[296] Whitney Quesenbery. The five dimensions of usability. *Content and complexity: Information design in technical communication*, pages 81–102, 2003.

[297] Titi Rahayu and Muhammad Aransyah. Analysis of qris user experience using the user experience questionnaire (ueq) method. *Jurnal Teknologi Informasi Universitas Lambung Mangkurat (JTIULM)*, 8(1):31–38, 2023. doi: 10.20527/jtiulm.v8i1.149.

[298] Liana Razmerita. *Ontology-Based User Modeling*, page 635–664. 2007. ISBN 9780387370194. doi: 10.1007/978-0-387-37022-4_23.

[299] Ateeq Ur Rehman, Rizwan Ali Naqvi, Abdul Rehman, Anand Paul, Muhammad Tariq Sadiq, and Dildar Hussain. A trustworthy siot aware mechanism as an enabler for citizen services in smart cities. *Electronics*, 9(6):918, 2020. doi: 10.3390/electronics9060918.

[300] Christian P Renaudin, Bernard Barbier, Ricardo Roriz, Didier Revel, and Michel Amiel. Coronary arteries: new design for three-dimensional arterial phantoms. *Radiology*, 190(2): 579–582, February 1994. ISSN 0033-8419. doi: 10.1148/radiology.190.2.8284422.

[301] Yacine Rezgui. Text-based domain ontology building using tf-idf and metric clusters techniques. *The Knowledge Engineering Review*, 22(4):379–403, December 2007. ISSN 0269-8889. doi: 10.1017/S0269888907001130.

[302] Petar Ristoski and Heiko Paulheim. Semantic web in data mining and knowledge discovery: A comprehensive survey. *Journal of Web Semantics*, 36:1–22, January 2016. ISSN 15708268. doi: 10.1016/j.websem.2016.01.001.

[303] Jane Ritchie, Jane Lewis, Carol McNaughton Nicholls, Rachel Ormston, et al. *Qualitative research practice: A guide for social science students and researchers*. 2013.

[304] Y. Rogers, H. Sharp, and J. Preece. *Interaction design beyond human-computer interaction*. 2004. doi: 10.5555/58015.

[305] Stephanie Rosenbaum, Chauncey E Wilson, Timo Jokela, Janice A Rohn, Trixi B Smith, and Karel Vredenburg. Usability in practice: user experience lifecycle-evolution and revolution. In *CHI'02 extended abstracts on Human factors in computing systems*, pages 898–903, April 2002. ISBN 1581134541. doi: 10.1145/506443.506649.

[306] Susi Ross, Magnus Ramage, and Yvonne Rogers. Petra: Participatory evaluation through redesign and analysis. *Interacting with Computers*, 7(4):335–360, 1995.

[307] V. Roto, E. Law, A. Vermeeren, and J. Hoonhout. User experience white paper: bringing clarity to the concept of user experience. *Result from Dagstuhl Seminar on Demarcating User Experience, number 10302*, 2011. doi: 10.14236/ewic/HCI2011.2.

[308] Virpi Roto, Marianna Obrist, and Kaisa Vaananen-Vainio-Mattila. User experience evaluation methods in academic and industrial contexts. In *Proceedings of the Workshop UXEM*, volume 9, pages 1–5. Citeseer, 2009.

[309] Tanushree Roy, Amara Tariq, and Satadru Dey. A socio-technical approach for resilient connected transportation systems in smart cities. *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[310] Vanessa Russo. Digital economy and society index (desi). european guidelines and empirical applications. *Qualitative and Quantitative Models in Socio-Economic Systems and Social Work*, 208:427, 2019. doi: 10.1007/978-3-030-18593-0_31.

[311] Michele Ruta, Floriano Scioscia, Giuseppe Loseto, and Eugenio Di Sciascio. Semantic-based resource discovery and orchestration in home and building automation: A multi-agent approach. *IEEE Transactions on Industrial Informatics*, 10(1):730–741, February 2013. ISSN 1551-3203. doi: 10.1109/TII.2013.2273433.

[312] Richard M. Ryan and Edward L. Deci. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1):68, 2000. doi: 10.1037/0003-066x.55.1.68.

[313] Marta Sabou. An introduction to semantic web technologies. *Semantic Web Technologies for Intelligent Engineering Applications*, pages 53–81, 2016.

[314] Rafael Sacks, Chuck Eastman, Ghang Lee, and Paul Teicholz. *BIM Handbook: A Guide to Building Information Modeling for Owners, Designers, Engineers, Contractors, and Facility Managers*. August 2018. ISBN 9781119287537. doi: 10.1002/9781119287568.

[315] Mandeep Kaur Saggi and Sushma Jain. A survey towards an integration of big data analytics to big insights for value-creation. *Information Processing & Management*, February 2018. ISSN 03064573. doi: 10.1016/j.ipm.2018.01.010.

[316] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. Tweet analysis for real-time event detection and earthquake reporting system development. *IEEE Transactions on Knowledge and Data Engineering*, 25(4):919–931, 2012. doi: 10.1109/TKDE.2012.29.

[317] Dina Salah, Richard F. Paige, and Paul Cairns. Integrating agile development processes and user centred design- a place for usability maturity models? In *Lecture Notes in Computer Science*, 2014. doi: 10.1007/978-3-662-44811-3_7.

[318] Yoshiki Sameshima and Peter Kirstein. Authorization with security attributes and privilege delegation: Access control beyond the acl. *Computer Communications*, 20(5):376–384, 1997.

[319] E. B. N. Sanders and P. J. Stappers. Co-creation and the new landscapes of design. *Co-design*, 4(1):5–18, 2008. doi: 10.1080/15710880701875068.

[320] Graig Sauer, Jonathan Holman, Jonathan Lazar, Harry Hochheiser, and Jinjuan Feng. Accessible privacy and security: a universally usable human-interaction proof tool. *Universal Access in the Information Society*, 9(3):239–248, August 2010. ISSN 1615-5289. doi: 10.1007/s10209-009-0171-2.

[321] Mark Saunders, PHILIP Lewis, and ADRIAN Thornhill. Research methods. *Business Students 4th edition Pearson Education Limited, England*, 2007.

[322] Mark Saunders, Philip Lewis, and Adrian Thornhill. *Research methods for business students*. 5th edition, 2009.

[323] MN Saunders. Choosing research participants. *Qualitative organizational research: Core methods and current challenges*, pages 35–52, 2012. doi: 10.4135/9781526435620.n3.

[324] Kim Schouten and Flavius Frasincar. Survey on aspect-level sentiment analysis. *IEEE Transactions on Knowledge and Data Engineering*, 28(3):813–830, March 2015. ISSN 1041-4347. doi: 10.1109/TKDE.2015.2485209.

[325] G. Schuh, J. Kuntz, K. Heeg, P. Jussen, J. Koch, and S. Breunig. Identification of variant-creating factors in product service systems. In *2016 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, pages 582–586, December 2016. ISBN 978-1-5090-3665-3. doi: 10.1109/IEEM.2016.7797942.

[326] Ulrike Schultze and Edwin L Cox. *Investigating the contradictions in knowledge management*. 1998.

[327] Rossitza Setchi and Obokhai K. Asikhia. Exploring user experience with image schemas, sentiments, and semantics. *IEEE Transactions on Affective Computing*, 10(2):182–195, 2019. doi: 10.1109/TAFFC.2017.2705691.

[328] Nigel Shadbolt, Tim Berners-Lee, and Wendy Hall. The semantic web revisited. *IEEE intelligent systems*, 21(3):96–101, May 2006. ISSN 1541-1672. doi: 10.1109/MIS.2006.62.

[329] Abdulrahman Sharida, Allam Hamdan, and AL-Hashimi Mukhtar. Smart cities: The next urban evolution in delivering a better quality of life. In *Toward Social Internet of Things (SIoT): Enabling Technologies, Architectures and Applications*, pages 287–298. 2020. doi: 10.1007/978-3-030-24513-9_16.

[330] A. Sheth. Internet of things to smart iot through semantic, cognitive, and perceptual computing. *IEEE Intell. Syst*, 31(2):108–112, March 2016. ISSN 1541-1672. doi: 10.1109/MIS.2016.34.

[331] Maulshree Singh, Evert Fuenmayor, Eoin P Hinchy, Yuansong Qiao, Niall Murray, and Declan Devine. Digital twin: Origin to future. *Applied System Innovation*, 4(2):36, 2021.

[332] Neha Singh, Clark Hu, and Wesley S Roehl. Text mining a decade of progress in hospitality human resource management research: Identifying emerging thematic development. *International Journal of Hospitality Management*, 26(1):131–147, 2007. doi: 10.1016/j.ijhm.2005.10.002.

[333] Tarsem Singh. A comprehensive review of text mining. *International Journal of Computer Science and Information Technologies*, 7(1):167–169, 2016.

[334] Paul Slovic, Melissa L Finucane, Ellen Peters, and Donald G MacGregor. The affect heuristic. *European journal of operational research*, 177(3):1333–1352, March 2007. ISSN 03772217. doi: 10.1016/j.ejor.2005.04.006.

[335] Ian Sommerville and Gerald Kotonya. Requirements engineering: a roadmap. *Proceedings of the Conference on the Future of Software Engineering*, 2000. doi: 10.1145/336512.336523.

[336] Juyoung Song, Tae Min Song, and Jin Ree Lee. Stay alert: Forecasting the risks of sexting in korea using social big data. *Computers in Human Behavior*, 81:294–302, 2018. doi: 10.1016/j.chb.2017.12.035.

[337] Wenyan Song and Tomohiko Sakao. A customization-oriented framework for design of sustainable product/service system. *Journal of Cleaner Production*, 140:1672–1685, January 2017. ISSN 09596526. doi: 10.1016/j.jclepro.2016.09.111.

[338] Sergey Sosnovsky and Darina Dicheva. Ontological technologies for user modelling. *International Journal of Metadata, Semantics and Ontologies*, 5(1):32–71, 2010. ISSN 1744-2621. doi: 10.1504/IJMSO.2010.032649.

[339] ISO Standards. Iso 9241-210:2010, ergonomics of human-system interaction - part 210: Human-centered design for interactive systems, 2010.

[340] Rudi Studer, V Richard Benjamins, and Dieter Fensel. Knowledge engineering: principles and methods. *Data and Knowledge Engineering*, 25(1-2):161–197, 1998. doi: 10.1016/S0169-023X(97)00056-6.

[341] K. Su, J. Li, and H. Fu. Smart city and the applications. In *Electronics, Communications and Control (ICECC), 2011 International Conference on*, page 1028–1031. IEEE, September 2011. ISBN 978-1-4577-0320-1. doi: 10.1109/ICECC.2011.6066743.

[342] Mari Carmen Suarez-Figueroa, Asuncion Gomez-Perez, and Mariano Fernandez-Lopez. The neon methodology for ontology engineering. In Mari Carmen Suarez-Figueroa, Asuncion Gomez-Perez, Enrico Motta, and Aldo Gangemi, editors, *Ontology engineering in a networked world*, pages 9–34. 2012. ISBN 978-3-642-24793-4. doi: 10.1007/978-3-642-24794-1_2.

[343] B Subedi, Abeer Alsadoon, P W C Prasad, and A Elchouemi. *Secure paradigm for web application development*, page 1–6. September 2016. ISBN 978-1-5090-5398-8. doi: 10.1109/RoEduNet.2016.7753243.

[344] Bilal Succar. Building information modelling framework: A research and delivery foundation for industry stakeholders. *Automation in Construction*, 18(3):357–375, 2009. doi: 10.1016/j.autcon.2008.10.003.

[345] Bilal Succar and Willy Sher. A competency knowledge-base for bim learning. In *Australasian Journal of Construction Economics and Building-Conference Series*, volume 2, pages 1–10, 2014. doi: 10.5130/ajceb-cs.v2i2.3883.

[346] H. Sun. *Cross-cultural technology design: Creating culture-sensitive technology for local users*. 2012. doi: 10.1093/acprof:oso/9780199744763.001.0001.

[347] Jia Tan, Kari Ronkko, and Cigdem Gencel. A framework for software usability and user experience measurement in mobile industry. In *2013 Joint Conference of the 23rd International Workshop on Software Measurement and the 8th International Conference on Software Process and Product Measurement*, pages 156–164. IEEE, October 2013. ISBN 978-0-7695-5078-7. doi: 10.1109/IWSM-Mensura.2013.31.

[348] Fei Tao, He Zhang, Ang Liu, and Andrew Yc Nee. Digital twin in industry: State-of-the-art. *IEEE Transactions on Industrial Informatics*, 15(4):2405–2415, April 2018. ISSN 1551-3203. doi: 10.1109/TII.2018.2873186.

[349] Alenka Temeljotov Salaj and Bradley Loewen. The next generation of smart citizens: Experiences and inspiration from the+ cityxchange project. In Aleksandra Djukic, Aleksandra Krstic-Furundvzic, Eva Vanivsta Lazarevic, Milena Vukmirovic, Aleksandra Djukic, Aleksandra Krstic-Furundvzic, Eva Vanivsta Lazarevic, and Milena Vukmirovic, editors, *7th International academic conference on Places and Technologies*, volume 7 of *International academic conference on places and technologies*, pages 13–22. University of Belgrade, Faculty of Architecture Belgrade, 2020. ISBN 978-86-7924-240-2. doi: 10.18485/arh_pt.2020.7.ch1.

[350] Manfred Thuring and Sascha Mahlke. Usability, aesthetics and emotions in human–technology interaction. *International Journal of Psychology*, 42(4):253–264, August 2007. ISSN 0020-7594. doi: 10.1080/00207590701396674.

[351] Noam Tractinsky, Adi Katz, and D. Ikar. What is beautiful is usable. *Interacting With Computers*, 2000. doi: 10.1016/s0953-5438(00)00031-x.

[352] Brandon Truong, Cornelia Caragea, Anna Squicciarini, and Andrea H Tapia. Identifying valuable information from twitter during natural disasters. *Proceedings of the American Society for Information Science and Technology*, 51(1):1–4, 2014. doi: 10.1002/meet.2014.14505101162.

[353] Mikalai Tsytsarau and Themis Palpanas. Survey on mining subjective data on the web. *Data Mining and Knowledge Discovery*, 24(3):478–514, May 2012. ISSN 1384-5810. doi: 10.1007/s10618-011-0238-6.

[354] Alexandre N. Tuch, Javier A. Bargas-Avila, Klaus Opwis, and Frank H. Wilhelm. Visual complexity of websites: Effects on users' experience, physiology, performance, and memory. *International Journal of Human Computer Studies*, 67:703–715, 2009. ISSN 10715819. doi: 10.1016/j.ijhcs.2009.04.002.

[355] Eric J Tuegel, Anthony R Ingraffea, Thomas G Eason, and S Michael Spottswood. Reengineering aircraft structural life prediction using a digital twin. *International Journal of Aerospace Engineering*, 2011:1–14, 2011. ISSN 1687-5966. doi: 10.1155/2011/154798.

[356] Mark Turner, Barbara Kitchenham, Pearl Brereton, Stuart Charters, and David Budgen. Does the technology acceptance model predict actual use? a systematic literature review. *Information and Software Technology*, 52(5):463–479, 2010. doi: 10.1016/j.infsof.2009.11.005.

[357] K Udom. Building information modelling, 2012.

[358] R. Unger and C. Chandler. *A Project Guide to UX Design: For user experience designers in the field or in the making*. 2012.

[359] John Andrew van der Poll, Fisokuhle Hopewell Nyembe Nyembe, and Hugo Hendrik Lotriet. Formal methods for an agile scrum software development methodology. In *Proceedings of the International Conference on Advanced Technologies*, volume 11, pages 168–176, 2023. doi: 10.58190/icat.2023.35.

[360] Renaud Vanlande, Christophe Nicolle, and Christophe Cruz. Ifc and building lifecycle management. *Automation in construction*, 18(1):70–78, December 2008. ISSN 09265805. doi: 10.1016/j.autcon.2008.05.001.

[361] Hernan Vargas, Carlos Buil-Aranda, Aidan Hogan, and Claudia Lopez. A user interface for exploring and querying knowledge graphs (extended abstract). In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, pages 4785–4789, July 2020. ISBN 978-0-9992411-6-5. doi: 10.24963/ijcai.2020/666.

[362] Ravi Vatrapu and Daniel Suthers. Cultural considerations in computer supported collaborative learning. *Research and Practice in Technology Enhanced Learning*, 6(1):3–26, 2011.

[363] Viswanath Venkatesh. Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model. *Information Systems Research*, 11(4):342–365, 2000. doi: 10.1287/isre.11.4.342.11872.

[364] Viswanath Venkatesh, Michael G. Morris, Gordon B. Davis, and Fred D. Davis. User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3):425–478, 2003. doi: 10.2307/30036540. URL `https://doi.org/10.2307/30036540`.

[365] Arnold POS Vermeeren, Effie Lai-Chong Law, Virpi Roto, Marianna Obrist, Jettie Hoonhout, and Kaisa Vaananen-Vainio-Mattila. User experience evaluation methods: current state and development needs. In *Proceedings of the 6th Nordic conference on human-computer interaction: Extending boundaries*, pages 521–530. ACM, 2010. doi: 10.1145/1868914.1868973.

[366] Chad Vicknair, Michael Macias, Zhendong Zhao, Xiaofei Nan, Yixin Chen, and Dawn Wilkins. A comparison of a graph database and a relational database: a data provenance perspective. In *Proceedings of the 48th annual Southeast regional conference*, pages 1–6, April 2010. ISBN 9781450300643. doi: 10.1145/1900008.1900067.

[367] Iosif Viktoratos, Athanasios Tsadiras, and Nick Bassiliades. Providing a context-aware location based web service through semantics and user-defined rules. In *Proceedings of the 4th International Conference on Web Intelligence, Mining and Semantics (WIMS14) - WIMS '14*, pages 1–11, June 2014. ISBN 9781450325387. doi: 10.1145/2611040.2611083.

[368] Rebekka Volk, Julian Stengel, and Frank Schultmann. Building information modeling (bim) for existing buildings–literature review and future needs. *Automation in construction*, 38:109–127, March 2014. ISSN 09265805. doi: 10.1016/j.autcon.2013.10.023.

[369] K. Vredenburg, J. Mao, P. W. Smith, and T. Carey. A survey of user-centered design practice. *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves*, pages 471–478, 2002. doi: 10.1145/503376.503460.

[370] Ernestas Vysniauskas, Lina Nemuraite, and Algirdas Sukys. A hybrid approach for relating owl 2 ontologies and relational databases. In Peter Forbrig, Horst Gunther, Will van der Aalst, John Mylopoulos, Norman M. Sadeh, Michael J. Shaw, and Clemens Szyperski, editors, *International Conference on Business Informatics Research*, volume 64 of *Lecture notes in business information processing*, pages 86–101. Springer, 2010. ISBN 978-3-642-16100-1. doi: 10.1007/978-3-642-16101-8_8.

[371] Joni Wahyudi. The acceptance of a smartphone application for disaster: Technology acceptance model approach. *IOP Conference Series Earth and Environmental Science*, 1180(1):012002, 2023. doi: 10.1088/1755-1315/1180/1/012002.

[372] Karzan Wakil. Extensibility interaction flow modeling language metamodels to develop new web application concerns. *Kurdistan Journal for Applied Research*, 2(3), 2017. ISSN 24117684. doi: 10.24017/science.2017.3.23.

[373] Sheng-Ming Wang and Chieh-Ju Huang. User experience analysis on urban interaction and information service in smart city nodes. In *Proceedings of the Second International Symposium of Chinese CHI on - Chinese CHI '14*, pages 103–109, April 2014. ISBN 9781450328760. doi: 10.1145/2592235.2592251.

[374] Zhu Wang, Xingshe Zhou, Zhiwen Yu, Haipeng Wang, and Hongbo Ni. Quantitative evaluation of group user experience in smart spaces. *Cybernetics and Systems*, 41:105–122, 2010. ISSN 01969722. doi: 10.1080/01969720903584209.

[375] H. Wicaksono, P. Dobreva, P. Hafner, and S. Rogalski. Methodology to develop ontological building information model for energy management system in building operational phase. In Ana Fred, Jan L.G. Dietz, Kecheng Liu, and Joaquim Filipe, editors, *Knowledge Discovery, Knowledge Engineering and Knowledge Management*, volume 454 of *Communications in computer and information science*, page 168–181. 2013. ISBN 978-3-662-46548-6. doi: 10.1007/978-3-662-46549-3_11.

[376] Bernd W Wirtz, Vincent Gottel, Paul F Langer, and Marc-Julian Thomas. Antecedents and consequences of public administration's social media website attractiveness. *International Review of Administrative Sciences*, 86(1):38–61, 2020. doi: 10.1177/0020852318762310.

[377] D. Wixon, K. Holtzblatt, and S. Knox. Tracking real-time user experience (true) a comprehensive instrumentation solution for complex systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 443–452. ACM, 2008. doi: 10.1145/1357054.1357126.

[378] Peter Wright and John McCarthy. The value of the novel in designing for experience. In Antti Pirhonen, Pertti Saariluoma, Hannakaisa Isomaki, and Chris Roast, editors, *Future interaction design*, pages 9–30. 2005. ISBN 1-85233-791-5. doi: 10.1007/1-84628-089-3_2.

[379] P Wu and R Issa. Patterns and trends in building information modeling (bim) research: A latent semantic analysis. *International Journal of Construction Education and Research*, 11(4): 304–326, 2015. doi: 10.1080/15578771.2015.1058564.

[380] Wei Wu and Raja RA Issa. Key issues in workforce planning and adaptation strategies for bim implementation in construction industry. In *Construction Research Congress 2014: Construction in a Global Network*, pages 847–856, 2014. doi: 10.1061/9780784413517.087.

[381] Clarissa Castella Xavier and Marlo Souza. A basic approach for extracting and analyzing data from twitter. In Valter Roesler, Eduardo Barrere, and Roberto Willrich, editors, *Special Topics*

in Multimedia, IoT and Web Technologies, pages 185–211. 2020. ISBN 978-3-030-35101-4. doi: 10.1007/978-3-030-35102-1_7.

[382] Peipei Xia, Li Zhang, and Fanzhang Li. Learning similarity with cosine similarity ensemble. *Information Sciences*, 307:39–52, June 2015. ISSN 00200255. doi: 10.1016/j.ins.2015.02.024.

[383] Ruyue Xin, Hongyun Liu, Peng Chen, and Zhiming Zhao. Robust and accurate performance anomaly detection and prediction for cloud applications: a novel ensemble learning-based framework. *Journal of Cloud Computing*, 12(1), January 2023. ISSN 2192-113X. doi: 10.1186/s13677-022-00383-6.

[384] Li Ping Xu, Jian Hua Li, Ling Jiang Kong, Jun Wei Sun, and Hai Peng Wang. The relationship between urban planning and architectural design. *World Construction*, 7(1):13–15, January 2018. ISSN 23154586. doi: 10.26789/JWC.2018.01.004.

[385] Toshihiko Yamakami. Exploratory analysis of differences between social experience design and user experience design. In *16th International Conference on Advanced Communication Technology*, pages 769–773, February 2014. ISBN 978-89-968650-3-2. doi: 10.1109/ICACT.2014.6779066.

[386] J Yuan, Y Liu, and L Shen. Why sustainable construction? why not? an owner's perspective. *KSCE Journal of Civil Engineering*, 19(2):361–369, 2015. doi: 10.1007/s12205-015-0235-0.

[387] Yihong Yuan, Yongmei Lu, T Edwin Chow, Chao Ye, Abdullatif Alyaqout, and Yu Liu. The missing parts from social media–enabled smart cities: Who, where, when, and what? *Annals of the American Association of Geographers*, 110(2):462–475, 2020. doi: 10.1080/24694452.2019.1631144.

[388] B. Yuce and Y. Rezgui. An ann-ga semantic rule-based system to reduce the gap between predicted and actual energy consumption in buildings. *IEEE Transactions on Automation Science and Engineering*, 14(3):1351–1363, November 2015. doi: 10.1109/TASE.2015.2490141.

[389] Jian Zhang, Alessandro Simeone, Peihua Gu, and Bo Hong. Product features characterization and customerstextquoteright preferences prediction based on purchasing data. *CIRP Annals*, April 2018. ISSN 00078506. doi: 10.1016/j.cirp.2018.04.020.

[390] Liang Zhang, Jin Wen, Yanfei Li, Jianli Chen, Yunyang Ye, Yangyang Fu, and William Livingood. A review of machine learning in building load prediction. *Applied Energy*, 285:116452, March 2021. ISSN 03062619. doi: 10.1016/j.apenergy.2021.116452.

[391] Limao Zhang and Baabak Ashuri. Bim log mining: discovering social networks. *Automation in Construction*, 91:31–43, 2018. doi: 10.1016/j.autcon.2018.03.009.

[392] P. Zhang and G. M. Von Dran. An empirical study of online shopping customer satisfaction in china: a holistic perspective. *International Journal of Retail & Distribution Management*, 36 (11):919–940, 2008. doi: 10.1108/09590550810911683.

[393] Ping Zhang. Motivational affordances: Reasons for ict design and use. *Communications of the ACM*, 51(11):145–147, 2008.

[394] X Zhang, X Chen, J Zhang, and X Luo. Urbanization, economic growth and environmental pollution: Evidence from china. *Journal of Cleaner Production*, 182:849–858, 2018. doi: 10.1016/j.jclepro.2018.02.002.

[395] Y Zhang, X Wang, and Y Wang. Towards a semantic construction digital twin: Directions for future research. *Automation in Construction*, 105:102832, 2019. doi: 10.1016/j.autcon.2019.102832.

[396] Y Zhang, L Ding, and J Liu. Digital twin: Enabling technologies, challenges and open research. *Computer-Aided Civil and Infrastructure Engineering*, 35(2):154–168, 2020. doi: 10.1111/mice.12539.

[397] Xujuan Zhou, Xiaohui Tao, Jianming Yong, and Zhenyu Yang. Sentiment analysis on tweets for social events. In *Proceedings of the 2013 IEEE 17th international conference on computer supported cooperative work in design (CSCWD)*, pages 557–562. IEEE, 2013.

[398] Huiqi Zhu and Hanping Hou. Object unified identifier method in logistics resource integration. *Journal of Industrial Engineering and Management*, 8(1):184–202, March 2015. ISSN 2013-0953. doi: 10.3926/jiem.1357.

[399] Jun Zhu, Bill Chu, and Heather Lipford. *Detecting Privilege Escalation Attacks through Instrumenting Web Application Source Code*, page 73–80. June 2016. ISBN 9781450338028. doi: 10.1145/2914642.2914661.