A Real Time Urban Sustainability Assessment Framework For The Smart City Paradigm

Corentin Kuster

A thesis submitted in partial fulfilment of the requirement of the degree of Doctor of Philosophy

Cardiff School of Engineering
Cardiff University

2019

Acknowledgements

First and foremost, I would like to express my sincere appreciation to my supervisor, Professor Yacine Rezgui for accepting me into this program and the support he has given me across those four years. You have been of great guidance, helping me throughout my PhD journey and always present to answer to my doubts and interrogations.

I would also like to thank the Building Research Establishment (BRE) for giving me the opportunity to undertake this research project and giving me a platform to present my work to the industry.

A big "Merci" to my dear colleague Jean-Laurent Hippolyte for the many hours spent on projects, meetings, problem-solving, interns' co-supervision and all the simple chats about Britanny, La Reunion, good food and many more topics. People say that a PhD is a solitary experience but our collaborations have made it a really enjoyable experience. Also, thanks for the time you spent proof-reading about every manuscript I have written and for providing me with valuable comments on my work.

Thank to my always companion Gaila for her unlimited support and patience over those four years. You gave me the strength to accomplish this journey and I am fully grateful for that. Thanks to my family and friends back in France who have followed my progress from a distance and always believed in me.

Last but not least, thanks to all the incredible people I have met through those four years. So many faces, so many moments, so much fun. I am so grateful I have spent this journey with you all by my side. You definitely help in making it a life experience that won't be forgotten.

Declaration

STATEMENT 1 This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Signed ___________________________
Date ___________________________

STATEMENT 2
This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is it being submitted concurrently for any other degree or award (outside of any formal collaboration agreement between the University and a partner organisation)

Signed ___________________________
Date ___________________________

STATEMENT 3
I hereby give consent for my thesis, if accepted, to be available in the University's Open Access repository (or, where approved, to be available in the University's library and for inter-library loan), and for the title and summary to be made available to outside organisations, subject to the expiry of a University-approved bar on access if applicable.

Signed ___________________________
Date ___________________________

DECLARATION
This thesis is the result of my own independent work, except where otherwise stated, and the views expressed are my own. Other sources are acknowledged by explicit references. The thesis has not been edited by a third party beyond what is permitted by Cardiff University's Use of Third Party Editors by Research Degree Students Procedure.

Signed ___________________________
Date ___________________________

WORD COUNT ___64,573________
(Excluding summary, acknowledgements, declarations, contents pages, appendices, tables, diagrams and figures, references, bibliography, footnotes and endnotes)

Abstract

Cities have proven to be a great source of concerns on their impact on the world environment and ecosystem. The objective, in a context where environmental concerns are growing rapidly, is no longer to develop liveable cities but to develop sustainable and responsive cities.

This study investigates the currently available urban sustainability assessment (USA) schemes and outlines the main issues that the field is facing. After an extensive literature review, the author advocates for a scheme that would dynamically capture urban areas sustainability insights during their operation, a more user-centred and transparent scheme. The methodological approach has enabled the construction of a solid expertise on urban sustainability indicators, the essential role of the smart city and the Internet of Thing for a real-time key performance indicators determination and assessment, and technical and organisational challenges that such solution would encounter. Key domains such as sensing networks, remote sensing and GIS technologies, BIM technologies, Statistical databases and Open Governmental data platform, crowdsourcing and data mining that could support a real-time urban sustainability assessment have been studied.

Additionally, the use of semantic web technologies has been investigated as a mean to deal with sources heterogeneity from diverse data structures and their interoperability. An USA ontology has been designed, integrating existing ontologies such as SSN, ifcOWL, cityGML and geoSPARQL. A web application back-end has then been built around this ontology. The application backbone is an Ontology-Based Data Access where a Relational Database is mapped to the USA ontology, enabling to link sensors data to pieces of information on the urban environment.

Overall, this study has contributed to the body of knowledge by introducing an Ontology-Based Data Access (OBDA) approach to support real-time urban sustainability assessment leveraging sensors networks. It addresses both technical and organisational challenges that the smart systems domain is facing and is believed to be a valuable approach in the upcoming smart city paradigm.

The solution proposed to tackle the research questions still faces some limitations such as a limited validation of the USA scheme, the OBDA limited intelligence, an improvable BIM and cityGML models conversion to RDF or the lack of user interface. Future work should be carried out to overcome those limitations and to provide stakeholders a high-hand service.

Table of Contents

List of Figures

List of Tables

Abbreviations

AEC	Architecture, Engineering, Construction
AIM	Abductory Induction Mechanism
AMC	Amsterdam's Smart City
ANFIS	Adaptive Neuro-Fuzzy Inference Systems
ANN	Artificial Neural Network
ANOVA	Analysis Of Variance
API	Application Programming Interface
AR	Autoregressive
ARIMA	Autoregressive Integrated Moving Average
BEMS	Building Energy Management System
BIM	Building Information Model
BMS	Building Management System
BRE	Building Research Establishment
CAGR	Compound Annual Growth Rate
CBR	Case Based Reasoning
CHP	Combined Heat And Power
CN	Collaborative Network
COP	Coefficient Of Performance
CRF	Conditional Random Field
CUSP	Cardiff Urban Sustainability Platform
DHC	District Heating And Cooling
DL	Description Logic
DSS	Decision Support System
DUL	Dolce Ultra-Lite
EER	Energy Efficiency Ratio
ETSI	European Telecommunications Standards Institute
EU	European Union
FIR	Fuzzy Inductive Reasoning
FOAF	Friend Of A Friend
FOI	Feature Of Interest
GDPR	General Data Protection And Privacy Regulation
GHG	Greenhouse Gas
GIS	Geographic Information System
GML	Geography Markup Language
GPS	Global Positioning Systems
HTTP	Hypertext Transfer Protocol
HVAC	Heating, Ventilation And Air-Conditioning
ICT	Information And Communication Technologies
IFC	Industry Foundation Classes
IoT	Internet Of Things
IS	Information System
ISO	International Organization For Standardization

JDBC	Java Database Connectivity
KPI	Key Performance Indicator
LADAR	Laser Detection And Ranging
LANs	Local Area Networks
LiDAR	Light Detection And Ranging
LoD	Level Of Detail
M2M	Machine-To-Machine
MA	Moving Average
MAD	Median Absolute Deviation
MTR	M5’ Model Tree
NDVI	Normalized Difference Vegetation Index
O&M	Operation & Maintenance
OBDA	Ontology-Based Data Access
OGC	Open Geospatial Consortium
ORS	Ontology Requirements Specification
OWL	Web Ontology Language
QoS	Quality Of Service
QUDT	Quantity, Unit, Dimension And Type
RDBMS	Relational Database Management Systems
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
REGEX	Regular Expression
RF	Random Forest
RFID	Radio-Frequency Identification
RIF	Rule Interchange Format
RML	RDF Mapping Language
RMSE	Root Mean Squared Errors
ROI	Return On Investment
RQ	Research Question
SAREF	Smart Appliances Reference Ontology
SARIMA	Seasonal Autoregressive Integrated Moving Average
SEER	Seasonal Energy Efficiency Ratio
SOS	Sensor Observation Service
SOSA	Sensor, Observation, Sample, And Actuator
SQL	Structured Query Language
SSN	Semantic Sensor Network
SVM	Support Vector Machine
SWRL	Semantic Web Rule Language
TF	Tree Forest
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
USA	Urban Sustainability Assessment
USADB	Urban Sustainability Assessment Database
VE/VO	Virtual Organisation And Virtual Enterprises
VOC	Volatile Organic Compound
W3C	World Wide Web Consortium

WAN	Wide Area Networks
WDQS	Wikidata Query Service
WKT	Well Known Text
WWW	World Wide Web
XML	Extensible Mark-Up Language

Introduction

1.1 BACKGROUND

Globalisation has led to the creation of attractive poles that gather resources, key infrastructures and jobs [1]. This paradigm resulted in migration from rural to urban zones and cities and a sharp increase in the urban population. A city is defined by its functional characteristics, for example the coexistence of industrial, commercial, and political activities, and its aggregation of population [2]. Figure 1-1, taken from the World Urbanization Prospects 2018 website [3], gives an insight on the urban population by size class across the world in 2030. In Europe and North America, around 80% of the population lives in cities [3]–[5], a tendency that increases. The same goes for developing countries, especially Asia and Africa, booming economically, that are estimated to host two-thirds of their population in urban areas by 2050 [6].

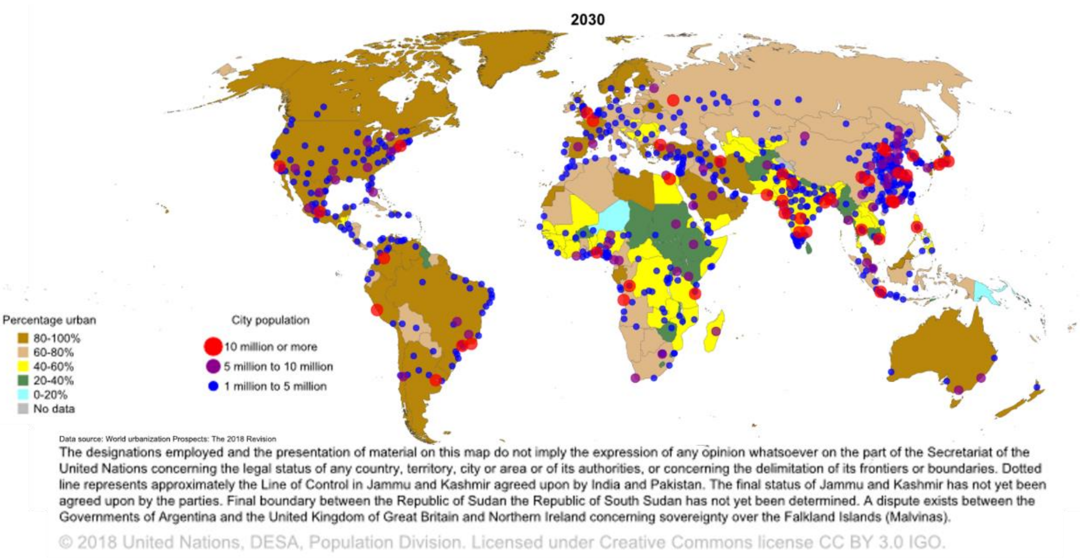


Figure 1-1 Percentage urban and urban agglomeration by size class (2018 United Nations, DESA, Population Division. Licensed under Creative Commons license).

This goes along a growing world population with an expected 2.5 billion increase in urban areas by 2050, 90% of which happening in Asia and Africa [5]. Figure 1-2, taken from the World Urbanization Prospects 2018 website [3], gives the urban agglomeration growth rates by size class across the world between 2018 and 2030. While Europe and North America have a growth rate <1%, exceeding this value in few cases, Asia and Africa show an urban growth rate of from 1 to 5%, going above 5% in some urban agglomerations.

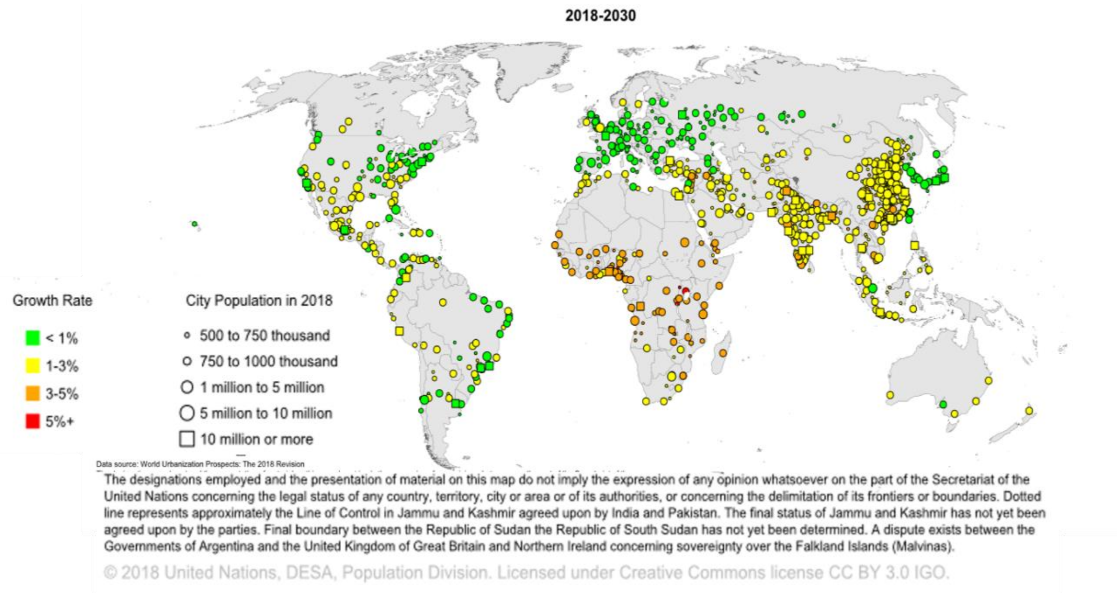


Figure 1-2 Growth rates of urban agglomerations by size class (2018 United Nations, DESA, Population Division. Licensed under Creative Commons license).

Those figures have triggered a real concern to governmental institutions that, in the few past decades, have increasingly included urban sustainability in their agenda for development. Indeed, urban areas are most likely to face challenges from traffic congestion, housing shortage, crowded streets to waste management, pollution, social conflicts etc [7]. In responses to those challenges, a set of sustainability schemes have been developed with the aim to holistically assess and tackle the major issues that cities are facing. Those schemes were firstly focused on new building development with frameworks such as BREEAM [8], LEED [9], CASBEE [10] , HQE [11] or DGNB [12]. However, urban areas are much more than the sum of their parts and a growing interest in regarding urban systems as holistic systems has emerged [13], [14]. New frameworks have been developed such as LEED-ND [15], BREEAM Communities [16] or CASBEE-UD [17] for instance, taking on a more global approach on urban areas. There are now a broad variety of schemes developed to fit particular locations, pressuring specific issues or adopting different approaches. Overall, a set of key performance indicators can be outlined from all those initiatives, defining, to some extent, what sustainability at the urban level is.

In parallel, cities are facing a technological revolution with an urge to become more intelligent and to become smart cities. Kitchin defines smart cities as:

> "*cities that, on the one hand, are increasingly composed of and monitored by pervasive and ubiquitous computing and, on the other, whose economy and*

governance is being driven by innovation, creativity and entrepreneurship, enacted by smart people."[18]

Smart cities are emerging, integrating information and communication technologies with remote sensing and actuation, and intelligence with machine learning to support decision making [18], [19]. A big part of this revolution lies on the Internet Of Things (IoT) and Big Data with the promise to interoperate and to catch anything, anywhere, at any time [20]. Multiples applications and services have already been implemented from smart homes and transports to smart energy grids and more [21]. In a context where urban sustainability is of prime concern, smart cities are believed to enhance a consistent view and understanding of city mechanisms, ultimately leading to more efficient decision making and management [18], [22]. Additionally, smart cities give the opportunity to assess sustainability in real or near real-time, missing characteristic of the current urban sustainability assessment (USA) schemes mentioned above. They provide a concrete advantage in the pursuit of operational and managerial actions for sustainability.

Overall, despite the promising picture of smart cities and IoTs supporting real-time urban sustainability assessment, leveraging data resources remains problematic leading to "*drowning in data*" [23]. Therefore, this thesis proposes a new real-time urban sustainability assessment framework centred on operational and managerial efforts that will fully leverage IoT potential via the integration of semantic web technologies to reach interoperability and enhance the system intelligence. Following this introduction, an extended literature review will set a strong theoretical background on urban sustainability, smart cities and semantic technologies. This is followed by the methodology employed in this research with design sciences and participatory action research that integrates stakeholders' contributions, real life case studies and smart city projects as sources of knowledge.

1.2 PROBLEM STATEMENT

The creation of a real-time urban sustainability assessment able to assist decision making on operational and managerial initiatives opens various conceptual and technical challenges.

Conceptually, current schemes sometimes diverge on their KPIs definition which demonstrates a lack of consensus around the domain of urban sustainability [24], [25]. Moreover, they are often implemented by experts in an ad-hoc way, at certain moments of urban area's development and operation, with little or no room for dynamism and adaptability in a complex and continually changing urban ecosystem [26]. Therefore, there is room for assessments, served by the IoT, which run in real or near real-time to help quick decision making during the operative stage of an urban system life cycle. The implementation of urban sustainability assessment in real-time questions what defines sustainability and especially sustainability within a real-time time frame. Such an approach will therefore open on the meaning of "real-time" for individual indicators and on how they can be measured on such time frame.

Technical challenges are inherent to the smart city paradigm and questions how to deploy holistic, ubiquitous services. Indeed, such an approach relies on heterogeneous information and systems that are asked to collaborate in order to provide value-added services [27]. Therefore, complete interoperability is key to leverage the full potential of the IoT [27]–[29]. If machine-to-machine communication procedures are well-defined with standardised gateways and communication protocols (Internet Protocol), at the application layer, interoperability remains an issue [30]. This issue can be tackled via the use of semantic web technologies that will semantically enhance information for greater machine interpretation [31]. The advantage is twofold: it set grounding for a common understanding of the information by heterogeneous systems, easing interoperability; and it creates linked data and enables to build intelligent applications.

Overall, the main problem is how to bring the urban sustainability assessment forward, using the smart city paradigm in order to deliver the best decision making support in urban areas in operation.

1.3 RESEARCH QUESTIONS

Based on the above discussion, the research core contribution focuses on answering the following questions:

RQ1. What are the issues that face the different stakeholders of an urban system toward sustainability assessment?

RQ2. How an effective urban sustainability assessment can help different parties of a city in their decision making?

RQ3. How can sustainability assessment leverage the smart city paradigm, specifically ICTs and the IoT?

RQ4. How can semantic web technologies unify heterogeneous data resources for holistic services and applications?

RQ5. How technological, human and financial assets relate to such service provision approach in the smart city paradigm?

1.4 RESEARCH DESCRIPTION

The present research adopted a pragmatic philosophy suggesting that the philosophical perspective should be tailored to the research questions [32]. A mixed approach is considered that gather both deductive and inductive approaches to harness the better of the two approaches. In that context, several strategies have been adopted including extended systematic reviews, structure and unstructured surveys and interviews, case studies, participatory action research and design sciences research. The research gap is first drawn by a systematic literature review of the current urban sustainability assessment schemes. The intensive literature review gives pieces of information on how the research gap can be filled, investigating the smart city paradigm, IoT and semantic web technologies. The knowledge base is then strengthened via participatory action research on projects developing smart solutions for district level energy provision. Those projects have provided case studies, expert opinions (via structure surveys such as DELPHI or informal exchanges) and technical assets, essential in the pursuit of the research questions. Finally, design science research, often employed in Information System, has been adopted for the development of the USA application. This iterative strategy allows the construction of solid expertise and meta-design and a clear understanding of the system assets and limitations. It takes full advantage of present case studies to develop and validate the application.

The objective of this research is to highlight current gaps in urban sustainability assessment schemes, to investigate valuable means to improve the issues in the field and to showcase the value of the proposed solution via its implementation in different case studies.

1.5 THESIS STRUCTURE

This thesis has been divided into 7 different chapters. Figure 1-3 gives the thesis structure with the 7 chapters along with some brief pieces of information on the research questions addressed in each of them.

The following chapters will develop on a new urban sustainability scheme suitable for real-time assessment, semantic web technologies and the validation of the conceptual framework via case studies and tests. Thoughts and interpretations on the feasibility and improvement of such system will be synthesized in the last chapter to conclude on the future of the field.

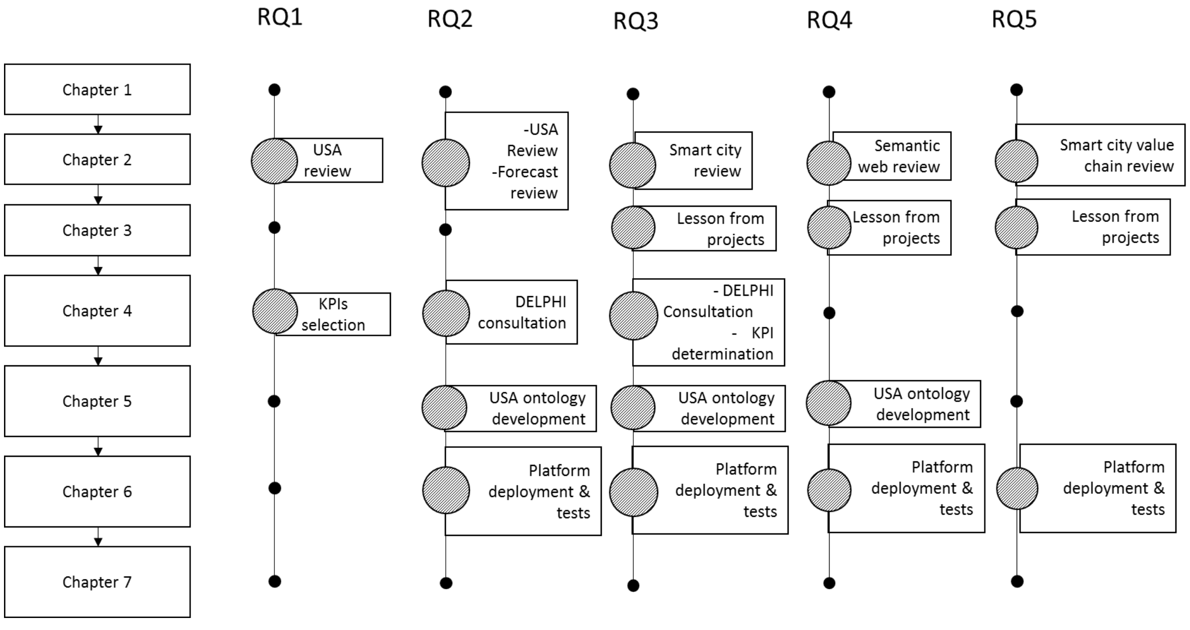


Figure 1-3 Thesis structure

1.5.1 Chapter Two: Literature Review

The second chapter gives an extended review on the major aspects and technologies involved in this research including (a) urban sustainability assessment frameworks, (b) internet of things & smart cities, (c) semantic technologies, (d) the smart city value chain and (e) decision making and the importance of forecasts.

1.5.2 Chapter Three: Design and Methodology

Chapter 3 introduces the methodology adopted in this research to address the research questions. The philosophical stance, approaches and strategies employed are described in details giving insights on how knowledge has been gathered. Adopting the pragmatic philosophy, the chapter presents the mixed approaches adopted including the participatory action research with the description of the various research projects involved, the collection of data via a DELPHI consultation and the case studies used. Additionally, a details description of the methodology for ontologies development is given in this chapter.

1.5.3 Chapter Four: Urban Sustainability Assessment Framework for Real-Time Assessment

This chapter explores the gap outlined in the literature review toward current urban sustainability assessment and presents a new framework as a synthesis of all of them. A total of 8 themes, 25 subthemes, 90 criteria and 193 indicators have been selected as the most encountered ones. For each indicator, means of measurement in real-time have been investigated centred around 5 main technologies namely (a) sensors, (b) remote sensing such as satellite or aerial high resolution imagery and GIS, (c) statistical databases and open government data, (d) BIM, and (e) crowdsourcing and data mining. When necessary, the author has provided academic references that support the feasibility of the measurement of some indicators in real-time. Finally, the DELPHI consultation on the newly developed scheme is presented along with the analysis of the responses in order to validate it in a real-time frame. Respondents have given their opinion on the relevance, feasibility of real-time measurements, logs frequencies and forecasts horizon required in this framework.

1.5.4 Chapter Five: Ontological Approach for a Real-Time Urban Sustainability Assessment Research

Chapter five presents the semantic web technology developed to support the USA framework, the USA ontology. The ontology requirement specifications are first introduced detailing on the main goals, domains and the scope of the ontology as well as defining the different uses and users. This includes the development of competency questions and the investigation of existing ontologies that will serve as a basis for the USA ontology. Thereafter, the USA ontology is described module by module and tested against a set of queries.

1.5.5 Chapter Six: The Framework Testing: Web Service for Real-time Sustainability Assessment

Chapter six introduces the final USA framework validation via the implementation of a web service for real-time sustainability assessment. The Ontology-Based Data Access (OBDA) approach adopted for such application technology is bridging a relational database where sensors' data are stored, and an RDF store where additional information such as BIM or GIS are stored using mappings between equivalent entities. Such approach is justified in this chapter, giving answers on why such technology is suitable for the purpose of this study. Each artefact involved in the development of the application is detailed including KPIs calculation and forecasts used and how they integrate such platform. Good attention is given to describe ONTOP, a fairly new technology that allows the development of OBDA. The OBDA is then tested against a set of queries to validate the feasibility and performance of the application. Finally, the intended design of the web service is given including its architecture and user-friendly interfaces.

1.5.6 Chapter Seven: Discussion and Conclusion

This chapter closes this thesis with a discussion of the finding and a summary. It gives the main research findings and how they answer the research questions. Remaining limitations are discussed and open on recommendations on future work that needs to be carried out.

1.6 SUMMARY

This chapter gave an overview of the research topic. It first presented the scope of the study by introducing key background elements that helped in the contextualisation of the research. It then introduced the main research aim and objectives along with the different research questions developed. The research and the followed methodology were then briefly described. Finally, the thesis structure was given with a succinct summary of each chapter as well as how they addressed certain of the research questions formulated. The following chapter gives the literature review findings and presents the key aspects involved in the research.

Literature Review

The literature review section contextualises the research by introducing the key fields drawn in the research questions. It allows the identification of the possible gaps that need to be filled and the considered technological solutions to those. First, a systematic review of the currently available urban sustainability assessment schemes has been done, identifying common gaps in the domain, especially the lack of operative application and dynamism. As a possible answer to these issues, the smart city paradigm and the Internet of Things (IoT) are introduced in great details along with the interoperability issues inherited by such perspective. The key contribution of semantic web technologies is then discussed as a mean to solve interoperability issues in smart cities and to shape ubiquitous linked data ecosystem. Smart city and urban sustainability-related ontologies are reviewed to illustrate the ongoing efforts in the domain. In a 4th section, the smart city value chain is discussed, illustrating the possible future smart cities' business model, the place of its actors and how the future dynamic urban sustainability assessment framework will fall within this perspective. Finally, the last section will open on the new conceptual approach induced by a dynamic urban sustainability assessment, as it may no longer be seen as a simple scoring scheme but rather as a decision support system for maintenance and operation. In this new vision, key business analytics are involved, especially forecasting models that have been reviewed.

2.1 URBAN SUSTAINABILITY ASSESSMENT FRAMEWORKS

This section begins with an overview of the field of sustainability assessment from building to urban scale. A systematic literature review of the current urban sustainability assessment schemes has been realised and the schemes have then been compared against their structure, sustainability dimension coverage, weighing and scoring systems. From that comparison, some gaps have been identified within the field, leaving room for possible improvement.

2.1.1 Overview of the Field

Sustainability at the neighbourhood scale is now considered essential for the development of sustainable cities [24]. As Jane Jacob already states in her book *"The Death and Life of Great American Cities"* [33]:

> *"A sustainable way of living should effortlessly derive from the way we design our sustainable neighbourhoods, as green neighbourhood developments are beneficial to the community and the individual as well as the environment".*

A good understanding of the neighbourhood structure and dynamic is crucial in order to achieve the city of tomorrow. Buildings can no longer be seen as individualities but as a part of a whole, interconnected, and ramified. An engineering approach toward a sustainable design of the neighbourhood considers the use of sustainable building assessment tools or frameworks.

Building sustainability assessment schemes are currently widely used to assess the environmental quality of a building. Many schemes exist such as BREEAM and its declinations (New construction [8], In-Use [34], Refurbishment [35]), LEED [9], CASBEE [10] , HQE [11] or DGNB [12]. Those are regularly employed by the AEC industry to certify the newly developed buildings as well as buildings in operation and refurbishment, ensuring a complete building life cycle evaluation and maintenance. If building sustainability assessments are well installed, the shift toward an urban scale sustainability assessment perspective is relatively new.

Upscaling sustainability assessment brings into the scope multiple new challenges that need to be considered. While a building is contained in its singular function, a neighbourhood is at the crossroad of various domains and its planning requires a much more holistic approach. Sustainability assessment at the urban level entails an intensive collect of information, flexibility as no generic model prevail, cross-scale and cross-disciplines considerations, an appropriate life cycle perspective and an increasingly influential human factor [36]. Therefore, a neighbourhood is much more than the sum of its parts that are buildings, and the fundamental question of sustainability at the urban level is, to some extent, different than at the building scale. Some even argue that building sustainability assessment lack of relevance as they minimise the crucial interaction between a building and its surrounding in term of space, function, social and cultural entity, aesthetic etc [36]–[39].

However, efforts have been made during the past decade regarding good practices in the design and operation of a neighbourhood [13], [14]. Companies and governmental institutions have released urban sustainability assessment schemes that aim at covering wider projects and perspectives [13], [40]. Table 2-1 shows the number of projects certified from 5 major sustainability assessment schemes.

Table 2-1 Sustainability schemes projects

	BREEAM	LEED	CASBEE	HQE	DGNB
Building scale	17048	84346	541	2409	2700
Urban scale	49	52	1	34	33
Source	[41] - From 2008 on	[42] - From 2008 on	[43] – As of July 2016	[44]	[45]

Looking at the figures in Table 2-1, it is quite clear that building certifications largely outnumber urban scale certifications. All urban scale certification schemes have been developed in the early 2010s which can explain the relatively few amounts of projects certified so far. It can be explained as well by the lower number of projects requiring certification at the urban scale. Overall, urban sustainability assessment is emerging at a slow pace. Nevertheless, urban scale sustainability assessment is growing in interest as all the major organisations leading the building sustainability assessment schemes have, in recent years, deployed efforts for their development. This not only witnesses the new trend emerging in the industry but also to a need from stakeholders for such type of assessment.

Those new frameworks constitute henceforth the basis for urban sustainability assessment and as such, many studies have been pursued to evaluate their performance and relevance. Haapio [46] has compared CASBEE-UD, BREEAM Communities and LEED-ND looking at the different indicators and criteria addressed as well as scoring systems they employ. In the same way, Orova [47] has realised a quantitative and qualitative analysis of each theme addressed in BREEAM Co., LEED-ND, CASBEE UD as well as DGNB Urban District. Ameen [14] and Gil [13] has studied in their respective research, the extent of sustainability coverage of BREEAM Co., LEED-ND, CASBEE-UD, SBToolPT-UP, Pearl Co. and GSAS/QSAS for the former and Citycad DPL, EcoCity Index, LEED-ND, Seeda SIC SN, Solutions, Spear and SUL for the latter. Finally, similar work has been undertaken by Sullivan [25] and Sharifi [24], [48]. All those studies overall aimed at drawing the current strengths and weaknesses of the available frameworks and at providing recommendations for a better scheme.

2.1.2 Frameworks analysis

The past decade has seen a sharp increase in urban sustainability assessment frameworks across the world. They give a hint on the definition of sustainability at the urban scale, helping practitioners to better understand its complex environment. Therefore, a systematic review of the existing USA frameworks will substantially improve the knowledge base and will allow spotting potential gaps that need to be fulfilled. They constitute a solid base for the construction of a new framework.

In the pursuit of a systematic literature review, one must first define search criteria related to the original research questions. In the present case, the literature material must contain "*technical specification"* of "*existing sustainability assessment schemes*" at the "*urban level*" and "*internationally distributed*". Considerable efforts have been done by the company Criterion Planners, an American urban and regional planning firm, in the creation of a global registry of urban sustainability rating tools [40]. The online tool called "Transformative tool" [49] gathers 61 tools distributed in 21 different countries and classified in various function categories, namely "cities", "planned neighbourhoods", "existing neighbourhoods, "All neighbourhoods", "Landscapes and Parks", "Transportation and Infrastructure" and "Special Purposes". All those frameworks have been developed by governmental institutions or private organisations and are currently used for profit or non-profit purposes in the evaluation of urban environments. Consequently, this registry has been used to explore the existing USA instances and has served as a basis for the systematic review since it complies with the search criteria for literature materials.

All the 61 frameworks have been primly analysed based on their specific purposes, the themes they addressed, their location, the year they have been launched, their physical scale etc. Out of the 61 schemes present in the registry, 32 appeared to be irrelevant because of a lack of exploitable information and/or a too specific purpose that does not cover holistically the domain of urban sustainability. Consequently, 29 USA frameworks remain in the scope of the study for an in-depth examination. Table 2-2 presents a brief overview of the 29 frameworks and their global specificities.

Table 2-2 USA frameworks [40]

Tool Name	Provider Name	Launch Year	Base Country	Available Countries	Scale	Reference
BERDE for Clustered Residential Development	Philippine Green Building Council	2013	Philippines	Philippines	Neighbourhood	[50]
BREEAM Communities	BRE Global Ltd	2009	U.K.	Europe, parts of The Middle East and Africa	Neighbourhood	[16]
CASBEE for Cities	Institute for Building Environment & Energy Conservation	2011	Japan	Global	Entire Cities	[51]
CASBEE for Urban Development	Institute for Building Environment & Energy Conservation	2006	Japan	Japan	Neighbourhood	[17]
CEEQUAL for Projects	CEEQUAL Ltd	2004	U.K.	Global	Neighbourhood up to entire cities	[52]
Comprehensive Plans for Sustaining Places	American Planning Association	2014	U.S.	n/a	Entire Cities	[53]
DGNB for Urban Districts	German Sustainable Building Council	2011	Germany	Global	Neighbourhood	[54]
EcoDistricts Protocol	EcoDistricts	2016	U.S.	Global	Neighbourhood up to entire cities	[55], [56]
EcoQuartiers	Ministry of Housing, Equality, and Rural Policy	2009	France	France	Neighbourhood	[57]
ELITE Cities	Ministry of Environmental Protection, China	2012	China/U.S.	China	Entire Cities	[58]
Enterprise Green Communities	Enterprise Partners	2011	U.S.	U.S.	n/a	[59]
EnviroDevelopment	Urban Development Institute of Australia	2006	Australia	Australia	Neighbourhood	[60]
Envision	Institute for Sustainable Infrastructure	2012	U.S.	North America	Neighbourhood up to entire cities	[61]
GBI Township Tool	Green Building Index	2011	Malaysia	Malaysia	Neighbourhood	[62]

Global Sustainability Assessment System for Districts	Gulf Organization for Research & Development	n/a	Qatar	Middle East	Neighbourhood	[63]
Gold Standard Cities	The Gold Standard	2015	Switzerland	Global	Entire Cities	[64]
Green Mark for Districts	Building & Construction Authority of Singapore	2009	Singapore	Singapore	Neighbourhood	[65]
Green Mark for Infrastructure	Building & Construction Authority of Singapore	2009	Singapore	Singapore	Neighbourhood up to entire cities	[66]
Green Star - Communities	Green Building Council of Australia	2012	Australia	Australia	Neighbourhood	[67]
HQE Urban Planning	Cerway/ Certivea/ Cerqual	2011	France	Global	Neighbourhood up to entire cities	[68]
IGBC Green Townships	Indian Green Building Council	2010	India	n/a	Neighbourhood up to entire cities	[69]
LEED for Neighborhood Development	U.S. Green Building Council	2009	U.S.	Global	Neighbourhood	[70]
Living Community Challenge	International Living Future Institute	2014	U.S.	Global	Neighbourhood up to entire cities	[71]
National Green Bldg Std for Land Development	Home Innovation Research Labs	2009	U.S.	U.S.	Neighbourhood	[72]
Pearl Rating System for Estidama - Community	Abu Dhabi Urban Planning Council	2010	Abu Dhabi	Abu Dhabi	Neighbourhood	[73]
PEER	Green Business Certification Inc.	2013	U.S.	Global	Neighbourhood up to entire cities	[74]
STAR Community	STAR Communities	2012	U.S.	North America	Entire Cities	[75]
Symbiocity	SKL International	2008	Sweden	Global	Entire Cities	[76]
Tool for Sustainable Urban Development	Realdania By	2007	Denmark	Denmark	Neighbourhood	[77]

Those 29 frameworks are now the basis for the study and creation of a new real-time urban sustainability assessment scheme. The different schemes structure, themes, criteria and indicators, underlying weighing system, features for local adaptability and sustainability dimension coverage are described and compared in the following sections, giving a global picture to the state of the field.

2.1.2.1 *Framework structure*

In a large majority of the cases, the assessment frameworks are structured in the hierarchical theme-criteria-indicator pattern where indicators are included into some criteria and criteria included into some themes. Sharifi [24] defines themes as "*broad topics of concern to sustainability*" such as climate and resources, or urban design for example. In term of abstraction, they must be wide enough to include various aspects and dimensions of sustainability but narrow enough to confine the focus on a particular matter. As Munier puts it [78], criteria are "*parameters used to evaluate the contribution of a project to meet the required objective*". They are much more tangible compared to themes and are linked to concrete objectives that need to be fulfilled in order to achieve the sustainability of an urban project. Efficient energy uses, water quality or a transport network penetration are few examples of criteria found in the literature. At the lowest level of the schemes' structure hierarchy are the indicators. They are described as qualitative or quantitative measurable elements [24], [79]. They are the building blocks of the sustainability scheme and are defined in an objective manner so that no interpretations are possible. For each indicator, a number of credits are allocated accordingly to its performance compared to fixed references. A sample of the theme-criteria-indicator structure of LEED-ND is shown in Figure 2-1 as a good example of the features described previously.

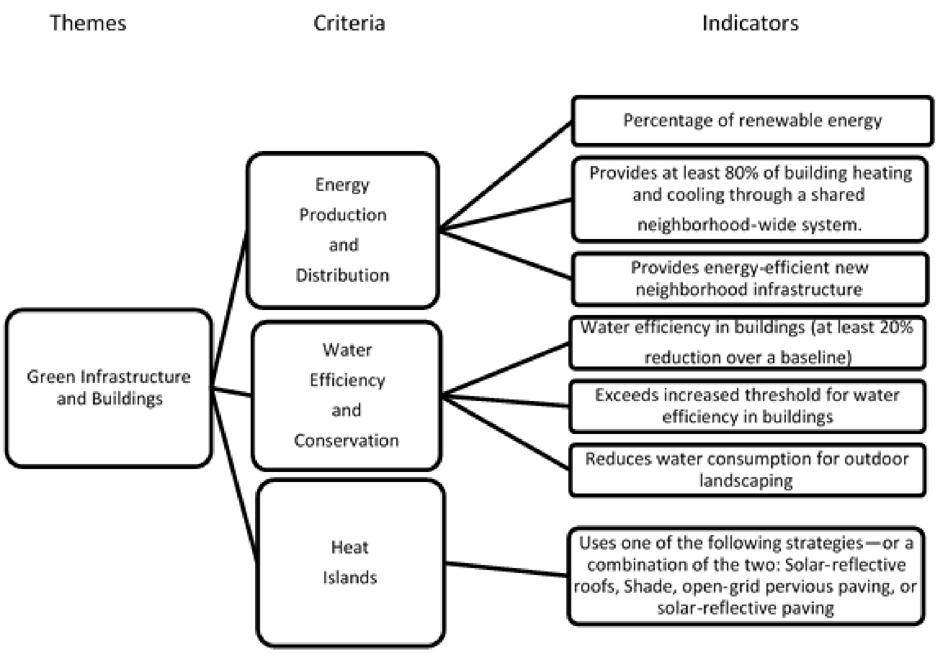


Figure 2-1 Sample of LEED-ND assessment scheme

Great differences can be noticed in the number of themes addressed in each of the 29 reviewed frameworks. Figure 2-2 shows the number of themes and criteria occurrences within each USA schemes. The number of themes varies from 4 for CASBEE for Cities, EcoQuartier or Gold Standards Cities up to 11 for BERDE for Clustered Residential Development with a median at 6 themes per schemes and an interquartile range at 3. In the high majority of the frameworks, global aspects such as "energy", "water", "transport" or "health" are considered as themes. However, a few frameworks use the sustainability dimension, namely "environmental", "social" and "economic", as some of their main themes; such as in CASBEE for cities, Gold standard Cities or DGNB-UD. Additionally, some aspects are defined as themes in a framework and as criteria in another. For example, the theme "Transportation" presents in ELITE cities or BERDE for Clustered Residential Development is referred to as a criterion in STAR Communities or BREEAM Communities. Same goes for criteria and indicator where in some cases, criteria are assimilated to indicators. Those disparities witness a lack of common understanding of the degree of abstraction that the themes and criteria must grasp.

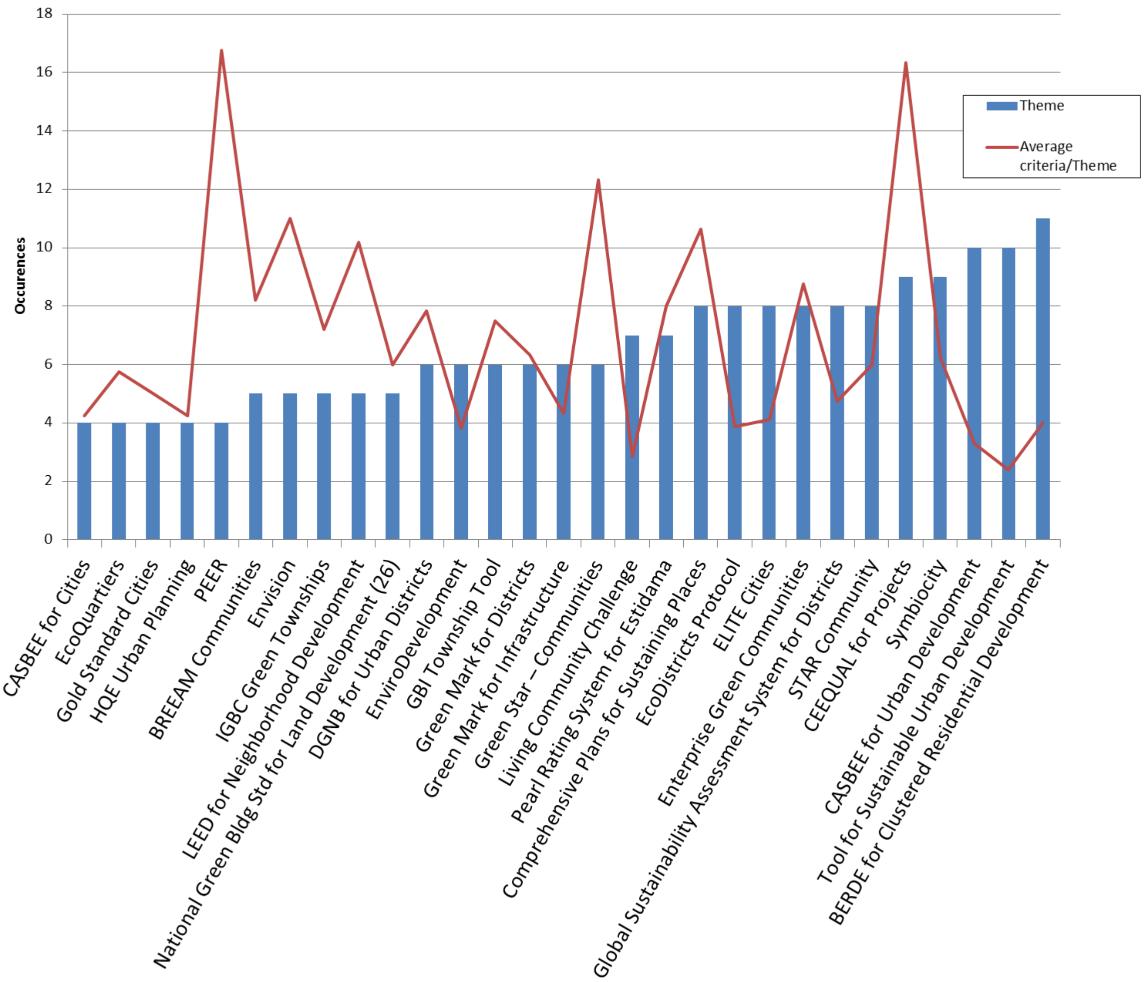


Figure 2-2 USA Frameworks themes and criteria occurrences

The differences between the frameworks are even more noticeable when looking at the number of criteria addressed in each of them. They vary from around 20 criteria in total in frameworks such as HQE, CASBEE for cities or Living community challenge up to 85 for Comprehensive Plans for Sustaining Places, with the particular case of CEEQUAL for Projects that defined 147 criteria. The median number of criteria per scheme is set at 38 with an interquartile range of 29. The number of criteria per theme goes from 1, for instance, “water consumption” is the unique criteria of the themes “Water” in Global Sustainability Assessment System for Districts, up to around 20. It exceeds 20 in really few cases such as in the “People and Communities” theme in CEEQUAL with 24 criteria or the “Reliability and Resilience” theme in PEER with 22 criteria. Figure 2-2 shows additionally the average number of criteria per theme in each scheme, the figures go from 2.40 criteria per theme in the case Tool for Sustainable Urban Development to 16 and 17 in the particular case of CEEQUAL and PEER. The median is set at 6 criteria per theme with an interquartile of 3.95.

Those differences between schemes can be explained by the different degree of abstraction given to each theme and criteria as stated above. Therefore, themes that cover a wider range of aspects will necessarily require a greater number of criteria. Additionally, some schemes voluntary emphasis on certain aspects and develop a greater amount of criteria for that particular focus.

Figure 2-3 shows the number of indicators addressed in each framework. Some frameworks did not fully provide the indicators they considered in their approach which explain the lack of data for some schemes in Figure 2-3. In the remaining frameworks, great disparities can be noticed. The number of indicators goes from 33 for Green Mark for Infrastructure and ELITE Cities up to 440 for Pearl Rating System for Estidama. Overall the median amount of indicator per framework is set at 116 indicators with an interquartile of 129. Relatively to the number of criteria, the number of indicators varies from 7.6 indicators per criteria for Green Mark for Infrastructure up to 55 for Pearl Rating System for Estidama with a median number at 14.5 and an interquartile at 7.8.

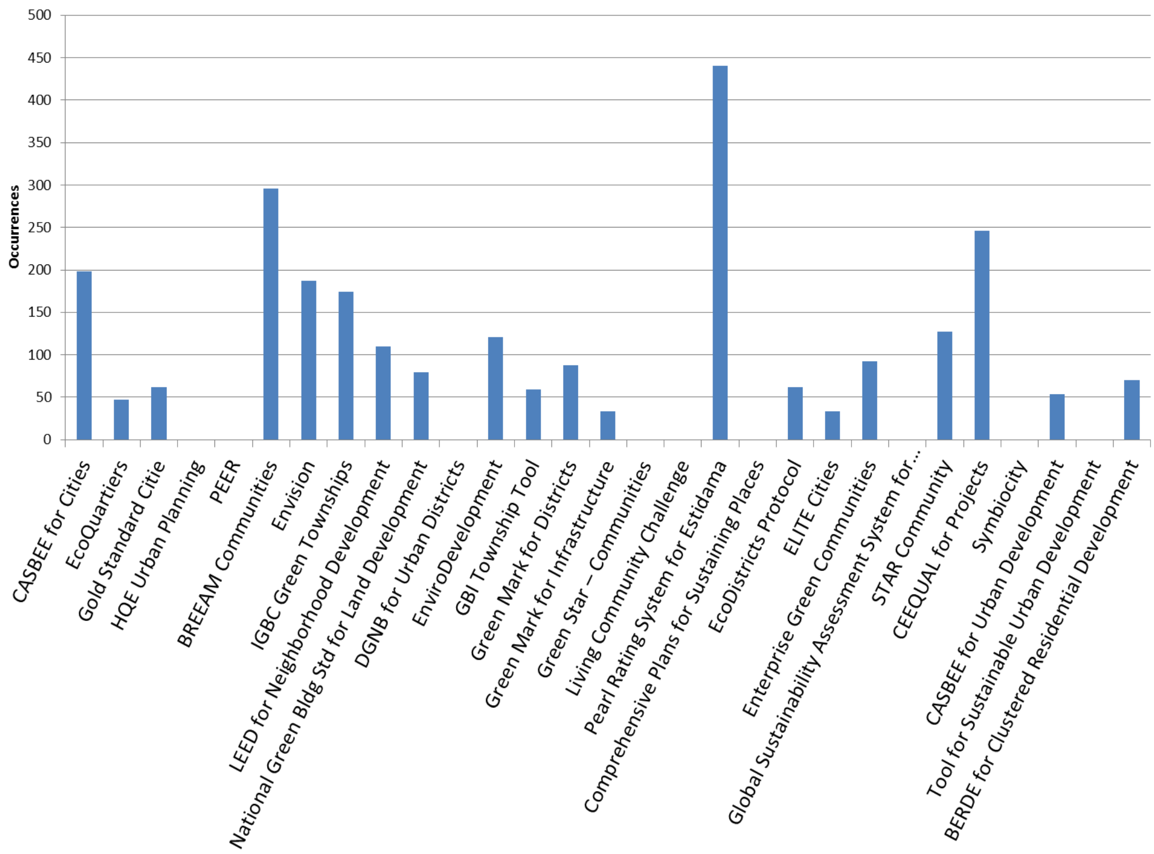


Figure 2-3 USA frameworks indicators frequency

Overall, a generic scheme model would comport approximatively 6 themes, each of which would include between 4 and 8 criteria that themselves would include between 11 and 18 indicators. Of course, this is only indicative and does not present a requirement concerning the structure that a USA scheme must have.

It is worth mentioning that 9 out of 29 of the frameworks did not present the indicators they considered in their approach. There is in a relatively high amount of cases a lack of transparency on this matter within the industry which ultimately lessens the understanding of a scheme and its mechanisms, and restricts constructive reflection around the topic of sustainability.

2.1.2.2 *Sustainability dimension*

The definition of sustainability is a delicate topic that has been subject to many debates since its first official introduction in the Brundtland Report [80]. The report defines sustainable development as a:

> *"development that meets the needs of the present without compromising the ability of future generations to meet their own needs. "*

If this common concept of sustainability is generally accepted, in the details it remains a controversial topic and there is still no real consensus toward what constitutes a sustainable development for cities. However, a well-accepted notion is the notion of the 3 pillars or dimensions of sustainability namely environmental, economic and social dimensions.

When looking at the different frameworks available for urban sustainability assessment it is interesting to compare the degree on which they cover those three pillars. Consequently, the indicators addressed in each framework and the weighing systems associated have been investigated. The repartition of credits and/or weighing coefficient of each indicators have been retrieve from the schemes and allocated to 8 main categories, namely "Resources and climate", "Land use and Ecology", "Urban Design", "Health and Well-Being", "Equity and Diversity", "Governance", "Innovation" and "Resilient Economy". "Resources and climate", "Land use and Ecology" cover Environmental aspects while "Urban Design", "Health and Well-Being", "Equity and Diversity" cover both Environmental and Social aspects and "Innovation", "Resilient Economy" the Economic aspects. Note that "Governance" introduces a 4th dimension, sometimes mentioned as a pillar of sustainability [25]. Figure 2-4 shows the repartition of every theme as a percentage of the total credits they hold considered in the 29 schemes. When a theme fits into one of the 8 predefined categories, it holds the colour of this one. In some occasions, themes cover more than one predefined categories. In that case, the themes can hold more than one colour corresponding to the several categories they cover. For that reason, Figure 2-4 does not present an exact repartition of the different themes in term of percentage but rather gives an approximated idea of it. Additionally, note that only 23 out of the 29 schemes have been taken for that study as some of them did not fully describe how the weighing system was linked to the themes addressed.

The theme "Resources and climate" is the most addressed, weighing around 29.7% of the total score median on the 29 schemes with an interquartile range of 27%. It is followed by the "Urban Design" and "Land Use and Ecology" with respectively 13.8% and 13.2% and an interquartile range of 17.2% and 9.9%. Comes next the "Health and well-being" counting for approximatively 10.1% with an interquartile range of 14.2%, "Governance" 9.3% and an interquartile range of 14.5%, "Resilient Economy" with 3.6% and 11.9% IQR and finally

“Innovation” and “Equity/Diversity” with 0% median and respectively 7% and 1.1% IQR. Overall, there is a clear focus on aspects such as “Resources and climate”, “Land Use and Ecology” and “Urban Design”; themes closely related to the environmental dimension of sustainability. In term of Social considerations, “Urban Design” and “Health and Well-Being” mostly covers this aspect with few regards to “Equity/Diversity”. “Innovation” presents some great disparities being not considered in most of the schemes but considered at 20 and 23% in 3 of them for instance. Same goes for “Governance” which is considered by all but in great differences. “Resilient Economy” is often present but with relatively low importance. It witnesses an unbalanced consideration regarding the different dimensions that define sustainability with a good emphasis on environmental aspects but disregarding socio-economic and cultural aspects. A good emphasis on environmental performance is observed and to some extent, social issue through the urban design, but there is a lack of concerns on socio-economic and cultural dimensions.

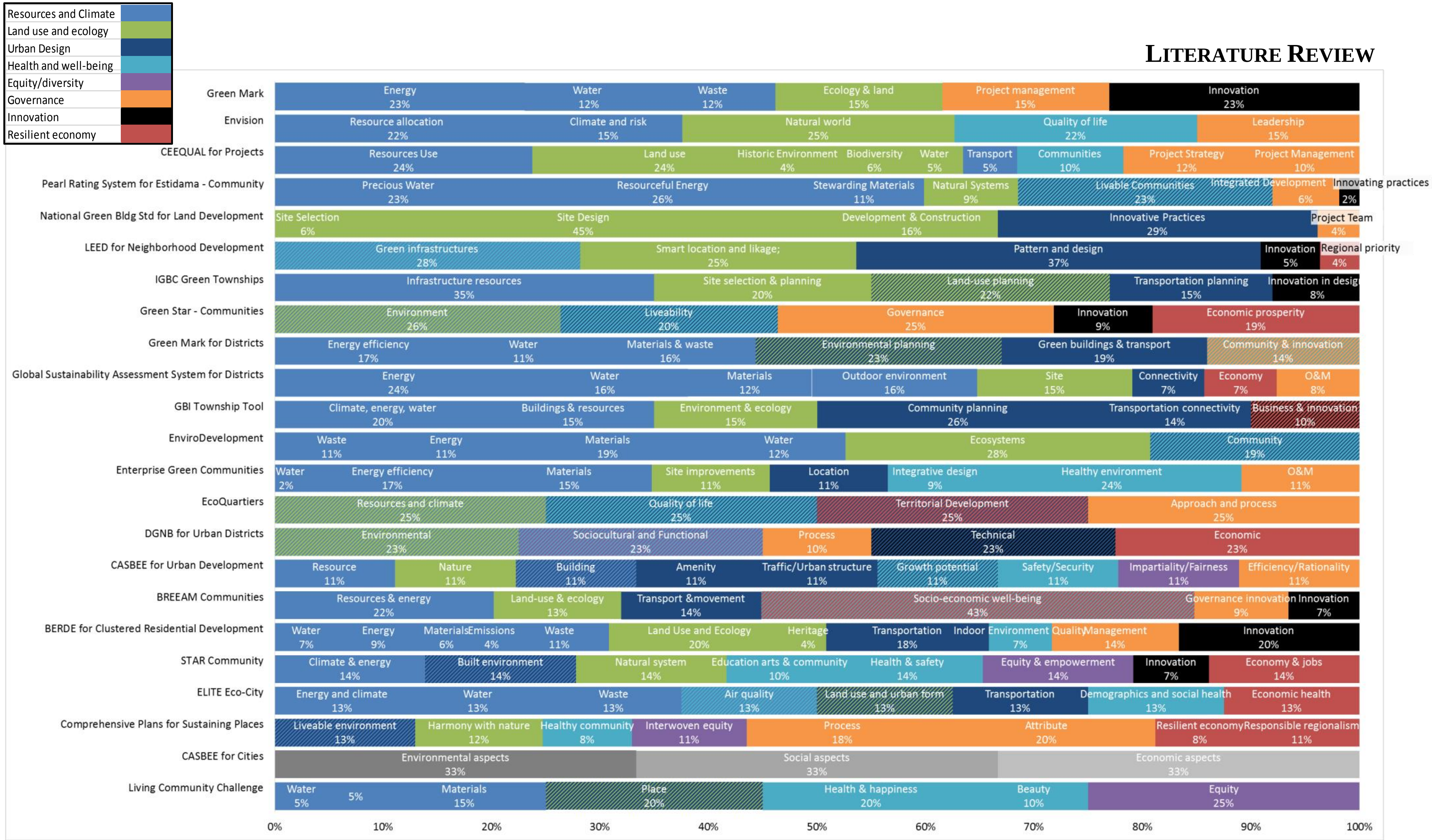


Figure 2-4 Themes repartition

2.1.2.3 *Weighing and Scoring Systems*

The schemes are underlaid by a weighing system that confers relative importance to each indicator. Table 2-3 presents the different weighing systems employed by each of the 29 USA frameworks. The system varies from a scheme to another. Indicator, criteria and themes can be weighted by a simple attribution of a number of credits and their sum like in LEED-ND, Green Mark for District or Pearl; or by the attribution of a weighing coefficient like in BREEAM Communities, DGNB for UD or CASBEE for UD. Additionally, some criteria or indicators that do not hold any credits or weight can often be found within the different frameworks. This is due to the fact that even if an indicator or criteria is considered within a scheme, this does not necessarily mean that it fulfils the necessary requirement to achieve sustainability [48]. Consequently, the schemes must have a system that ensures the minimum requirements. In their approach, the different framework considered that some criteria or indicators are mandatory and weighing them would be therefore unnecessary. Following the level in which the requirements are introduced within a scheme, they can be classified into 4 different types of weighing structure:

- *Requirements* based schemes that only considered requirements without any weighing system. In those cases, the scheme is more seen as a guide for good practices for sustainable development than a proper assessment (e.g. EcoDistricts Protocol, Living Community Challenge, Symbiocity).
- *Requirements with Bonuses* schemes that only considered requirements and introduced scoring as a final requirement of equal importance. In that case, there is a weighing system but the score does not prevail on the requirements and is more seen as a bonus (e.g. Enterprise Green Communities, EnviroDevelopment).
- *Score with prerequisites* schemes where requirements are prerequisites that need to be met to proceed to the assessment (e.g. BREEAM Communities, LEED-ND, DGNB for Urban Districts, and Pearl Rating System for Estidama – Community). A weighing system is present in that case on the non-mandatory criteria and/or indicators.
- *Score* based schemes that do not hold any mandatory requirements and are only based on the resulting score (e.g. STAR Community, Envision, Green Mark for Infrastructure). The entire assessment scheme is weighted in that case.

***Table 2-3 USA** frameworks weighing systems*

Tool	Rating structure	Comment
Ecodistricts Protocol	Requirements	The outcome is a clear set of project priorities to meet established performance goals
HQE Urban Planning	Requirements	Documentation audits are made at key stages of project development
Living Community Challenge	Requirements	All imperative is mandatory and no weighing system is applied. Thus KPIs have the same weigh by definition
CASBEE for Cities	Score	Criteria within themes have different weights but themes are equally weighted. Weights can change from a project to another.
Comprehensive Plans for Sustaining Places	Score	Set of practices scored according to the scoring matrix
ELITE Eco-City	Score	Indicators hold different weights in order to set the equal balance to primary categories. A second iteration of the method consists of the comparison with a similar town
Gold Standard Cities	Requirements with Bonuses	Some Principles, Activities, contextual and product requirements are applied. The SDGs are scored, quantifying the impact
STAR Community	Score	Apart from 2 themes, indicators hold different weights in order to set equal balance to primary categories.
Symbiocity	Requirements	An analysis of key issues is done at the earliest stage to define their importance. SWOT is used in this purpose
BERDE for Clustered Residential Development	Score with pre-requisites	Indicators are weighted by different amounts of points. Theme weight results from the sum of them. Innovation count for max 20 points included inside the other themes.
BREEAM Communities	Score with pre-requisites	Indicator ~~credits~~ have different weights. Theme weights result in the sum of them.
CASBEE for Urban Development	Score	Criteria within themes have different weights but themes are equally weighted. Weights can change from a project to another.
DGNB for Urban Districts	Score with pre-requisites	Apart for 1 theme, indicators hold different weigh in order to set equal balance to primary categories.
Ecoquartiers	Score	Criteria are individually considered and scored but no overall score is calculated.
Enterprise Green Communities	Requirements with Bonuses	All criteria are mandatory and 30 to 35 additionally points are required following the project type.
Envirodevelopment	Requirements with Bonuses	Most of the criteria are simply requirements while some will hold credits and a minimum score will serve as a requirement.
GBI Township Tool	Score with pre-requisites	Criteria are weighted by different amounts of points. Theme weight results from the sum of them. Requirements are mainly paperwork associated with the project management
Global Sustainability Assessment System for Districts	Score with pre-requisites	Criteria are weighted by different amounts of points. Theme weight results in the sum of them. Requirements are only submittal requirements.
Green Mark for Districts	Score with pre-requisites	Criteria are weighted by different amounts of points. Theme weight results in the sum of them. A few criteria are pre-requisites that do not hold any points.
Green Star - Communities	Score with pre-requisites	Criteria are weighted by different amounts of points. Theme weight result of the sum of them. Requirements are only submittal requirements.
IGBC Green Townships	Score with pre-requisites	Criteria are weighted by different amounts of points. Theme weight results in the sum of them. A few criteria are pre-requisites that do not hold any points.
LEED for Neighborhood Development	Score with pre-requisites	Criteria are weighted by different amounts of points. Theme weight result of the sum of them. A few criteria are pre-requisites that do not hold any points.

National Green Bldg Std for Land Development	Score with pre-requisites	Criteria are weighted by different amounts of points. Theme weight results from the sum of them. A few criteria are pre-requisites that do not hold any points.
Pearl Rating System for Estidama - Community	Score with pre-requisites	Criteria are weighted by different amounts of points. Theme weight result of the sum of them. A few criteria are pre-requisites that do not hold any points.
Tool for Sustainable Urban Development	Score	N/a
CEEQUAL for Projects	Score	Criteria are weighted by different amounts of points. Theme weight result of the sum of them. Not all criteria are mandatory.
Envision	Score	Criteria are weighted by different amounts of points. Theme weight results in the sum of them.
Green Mark for Infrastructure	Score	Criteria are weighted by different amounts of points. Theme weight results from the sum of them.
PEER	Score with pre-requisites	Criteria are weighted by different amounts of points. Theme weight results in the sum of them. The scheme distinguishes pre-requisite, core and bonuses criteria

2.1.3 Gaps in Current Urban Sustainability Assessments

2.1.3.1 *Lack of Consensus*

The analysis of the different urban sustainability assessment schemes has enabled the discovery of a certain number of issues and gaps. It is quite clear that no absolute model exists and that when it comes to evaluating the quality of a neighbourhood, each scheme applied what they considered as being of prime importance. Consequently, disparities exist between schemes which highlight a lack of common understanding of what defines sustainability. In the section 2.1.2.1 that presents the differences of structure in the 29 studied frameworks, the number of indicators and criteria addressed varies greatly from a scheme to another. Therefore, one could ask what is the appropriate number of indicator that needs to be considered to accurately evaluate sustainability. Tanguay et al. that have carried out 17 studies of the use of urban sustainable development indicators, came to the same conclusion concerning the disparities of the number of indicators [81]. They explain those variations by the nature of sustainability itself which is open to interpretation. Making an objective system, therefore, justifies the inclusions of a great number of indicators. On this matter, Ameen states that an increased number of criteria (and thus of indicators) can reduce overlaps between sustainability dimensions and themes, making it clearer for the different stakeholders [14]. Moreover, a well-known approach for indicator selection is the SMART approach namely:

> *"**S**pecific – target a specific area for improvement.*
> ***M**easurable – quantify or at least suggest an indicator of progress.*
> ***A**ssignable – specify who will do it.*
> ***R**ealistic – state what results can realistically be achieved, given available resources.*
> ***T**ime-related – specify when the result(s) can be achieved."*[82]

However, having too many indicators can result in a less generic model that is constrained to certain considerations. Consequently, the number of indicators must be carefully chosen in order to have a clear but reproducible model. This dichotomy explains the variations in the number of indicators between the different schemes.

In addition to the number of indicators addressed, the weighing system equally witnesses the lack of consensus toward the definition of sustainable development at the urban level. The weighing system assigns relative importance to the different aspects of sustainability. It is a subjective approach and therefore a controversial process. The differences in the weighing systems can mainly be explained by the various interests that the assessment provider has. Schemes will emphasis some aspects inherited by certain policies, institutions or individuals specific objectives [83]. Focusing on a certain domain also makes it possible to differentiate the scheme in an industry that is already encumbered and competitive. For instance, BREEAM Communities emphasis environmental concerns while LEED-ND seeks what is described as the "new urbanism" (site location and connectivity) [25]. Additionally, those differences are closely related to the local characteristics and how to best fit them [84]. Local adaptability is, therefore, a factor of change for the weighing system which is discussed later in 2.1.3.3.

Equally, Section 2.1.2.2 raised an important concern on the unequalled coverage of the various dimensions of sustainability. Indeed, the analysis of the 29 USA frameworks has shown that most of them greatly considered environmental (and its indirect economic impact) features at the expense of socio-economic features. Many studies that have reviewed current USA frameworks have observed the same outcome of unbalanced dimension considerations [25], [48], [85]. Those differences of treatment can be explained by the overall context in which those frameworks are embedded. As Davidson puts it, the neoliberal and technocratic approach of sustainability assessment has led the industry to emphasise technological fixes relative to economic outcomes [85]. Their engineering-driven approach focuses on aspects that can be improved and control by mean of technological assets with an indirect economic incentive. It disregards socio-economic aspects that are at the core of the social world in which cities are rooted.

Finally, the author has been confronted during this review to a lack of exploitable elements within the documentation available. If some frameworks describe in great details every indicator and the scoring behind them such as BREEAM Communities or LEED-ND, other omitted key elements that could have help in the full understanding of the schemes mechanism. This transparency issue has been spotted in multiple studies [13], [24], [25], [86], [87] and many advocates for a greater exposition of the core principle of USA frameworks. This would allow a more understandable and therefore inclusive approach to sustainability assessment.

Overall, those observations witness a lack of consensus in the domain on the definition of sustainability at the urban scale and how to assess it. This outcome is shared with other similar studies [14], [24], [25], [47]. There is a need for a constructive cooperation that would lead to the standardisation of the main features of urban sustainability assessment tools [81]. The domain would benefit from a common terminology and some principles that would constraint the selection of indicators (especially the mandatory ones) without affecting the flexibility of the assessment. It is more about defining a global frame than to impose strict proceedings. A good example to solve the issue would be the creation of a set of horizontal standardized methods, similar to CEN/TC 350 standards [88] which confers an integrated and holistic frame to the domain.

2.1.3.2 *Dynamism*

As Orova stated in her article “Comparison and evaluation of neighbourhood sustainability assessment systems” [47]:

> *“As the requirements towards neighbourhoods changes with different times, places, and cultures, the neighbourhood assessment systems have to constantly improve, and adapt to the current state of the environment.”*

This opens an interesting issue that is the issue of adaptability for sustainability assessments in term of temporal and local changes. Indeed, as environmental concerns rise among the expert community, requirements will become increasingly strict and narrow. Therefore, as requirements changes in time, there is a need for dynamic assessments in order to track the most recent changes. An assessment realised during the design phase of a project does not guaranty that it is going to comply with sustainable requirements over its entire lifespan. On the 29 frameworks studied, 24 primly focus on the design stage of a project, providing guidance over foreseen impacts of current practices in design. Four frameworks are specifically designed for the assessment of existing neighbourhoods or cities. And only one scheme (EcoDistricts Protocol) covers the operation of an existing district and the design stage. Note that this does not mean that all stages (design, construction, operation and post-occupancy) are not addressed in those frameworks but simply that they are addressed at one specific stage. This can be explained by the fact that most of those frameworks are destined to the construction and/or of the built environment industry that have development plans for an entire neighbourhood or more. Moreover, the assessment and maintenance of neighbourhood in operation is a more difficult task as it can encounter resistance from the residents and/or organisations [89], [90]. At the design stage of development, a dynamic scheme is not necessarily required as decisions are taken over long term aspects. Hence the predominance of "rigid" frameworks found in the literature and the lack of dynamism to catch changes over time.

Figure 2-5, adapted from "*Fourth Dimension in Building: Strategies for Avoiding Obsolescence*" [91] and the Pacific Northwest National Laboratory website [92], shows the effect of the good operation, maintenance and renewals on buildings' performance and lifespan. To some extent, since buildings are the building block of urban districts, this concept can be extrapolated to a neighbourhood. There are two essential aspects that can be learned from this figure: (1) as mentioned earlier, the minimum acceptable performance rises in time and this needs to be taken into account when assessing performance; (2) good operation, maintenance and periodic renewals have a dramatic influence on a building/neighbourhood performance and lifespan. Consequently, if most of the current frameworks are static and focused on long term impacts from the design stage, the field is missing dynamic schemes that are more specific to short term (up to real-time) changes. Those two complementary visions will ensure the development of robust projects in the long term as well as to keep up with increasing expectations via a good O&M.

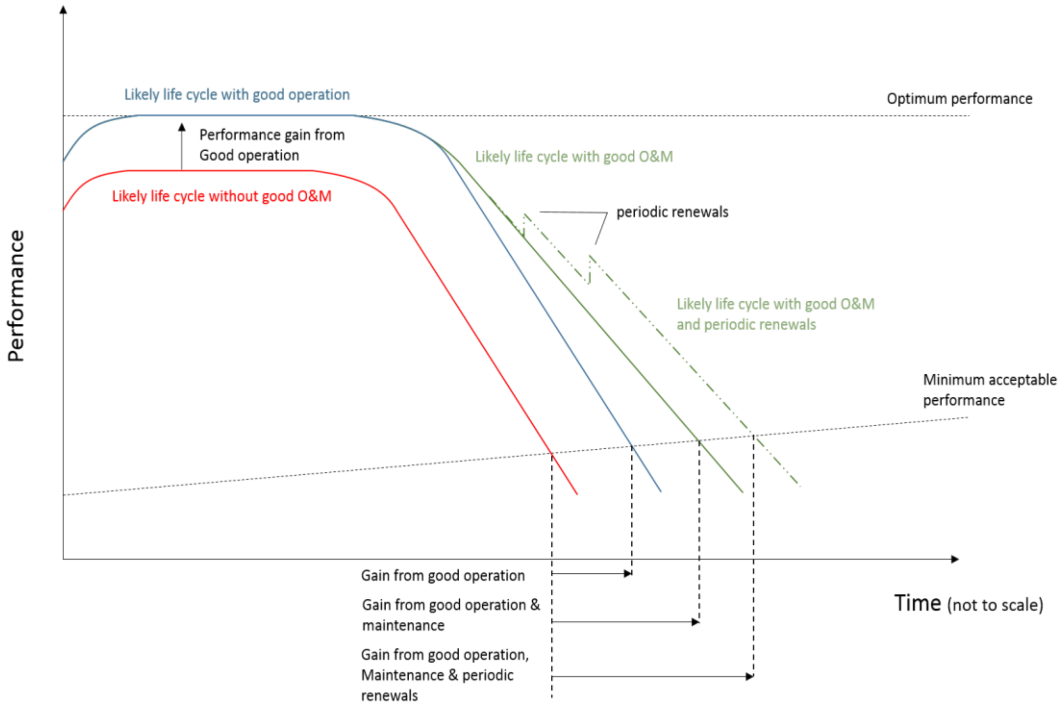


Figure 2-5 A conceptual view on building maintenance effect on performance

2.1.3.3 *Local adaptability and culture*

Orova statement on adaptability presents in the previous section also concerns the local adaptability of the sustainability assessment schemes. In the current landscape, most of urban districts sustainability assessments are suitable for their original country or region but miss flexibility in order to perform in other environments [46]. In fact, local adaptability is one of the major reasons for the frameworks disparities as stated earlier in section 2.1.3.1. Indeed, in their article [93], Chandratilake and Dias have compared the 'Energy Efficiency', 'Water Efficiency', and 'Site' criteria of 9 recognized assessment schemes applied in 9 different countries against the national annual CO2 emissions per capita, the percentage of land classified as water scarce and the population density of every one of those. They observed a clear correlation between the weighing system of certain criteria against aspects specific to certain locations, e.g. water stress in relation to weight for water. This emphasises the need for a specific framework to include processes for local adaptability. When it comes to this issue, mechanisms exist in some frameworks but are always on a case-by-case basis. For instance, EcoDistricts Protocol [55] and STAR Community [75] have integrated a consulting phase before assessment in order to redefine the weighing system so that it best fits a particular project. Same for CASBEE-Cities [51] that states that "*weighing coefficients may be adjusted to suit the specific circumstances of each* *city*". BREEAM Communities [16] have introduced an "International Bespoke" process to decide if the scheme suits for an international project and to allow occasional changes.

To conclude, if mechanisms exist so that the framework can be applied at different locations, those are rigid and lack of automatisms and flexibility. There is room in the field for processes that would ease local adaptability of neighbourhood sustainability assessment. Those changes must be made publicly available and transparent and follow a universalist code of ethic.

2.2 SMART CITIES, BIG DATA & THE INTERNET OF THINGS

In the present section, an overview of the emerging smart city paradigm is given. Core technological enablers are described in detail, from the IoT to BIM and GIS technologies as well as crowdsourcing and data mining. Those are believed to be the main contributors to big data, allowing business analytics and delivery of increasingly intelligent services. The investigated near-real-time urban sustainability assessment is therefore to be based on those core technologies, being fully integrated within the smart city vision.

2.2.1 Overview of the Field

Cities have become increasingly complex with a growing diversified population and a sprawling urban area. Historically, urban design was mainly focused on planning to accommodate citizens with little regard for environmental and social issues [94]. In contrast, during the past decade, substantial efforts have been made by experts in the study of the impacts of urbanisation on people and the environment. It is currently greatly recognized that the World is in the Anthropocene era, where human activity has a significant impact on the Earth’s geology and ecosystem [95]. In this prospect, a new paradigm has emerged for city stakeholders, where sustainability is at the core of the urbanisation thinking [96], [97]. In the meantime, those environmental challenges go along cutting-edge technologies in information and communication technologies. Those technologies enable sensing and in some cases control of environmental conditions leading to a better understanding of the city insights. Such an approach is leveraging sensing networks to overcome urban sustainability challenges. The incremental implementation of sensing and control devices triggers a new concept of the technology-driven city.

If many terms have been associated with cities that embraced ICT such as ‘Wired City’ [98], ‘Cyber city’ [99], ‘Digital City’ [100] , ‘Intelligent City’ [101] ,”Knowledge City” [102], the term “smart city” is clearly predominant since the early 2010s, as shown in Figure 2-6 , where the number of publications skyrocketed, which demonstrates the current interest in the topic.

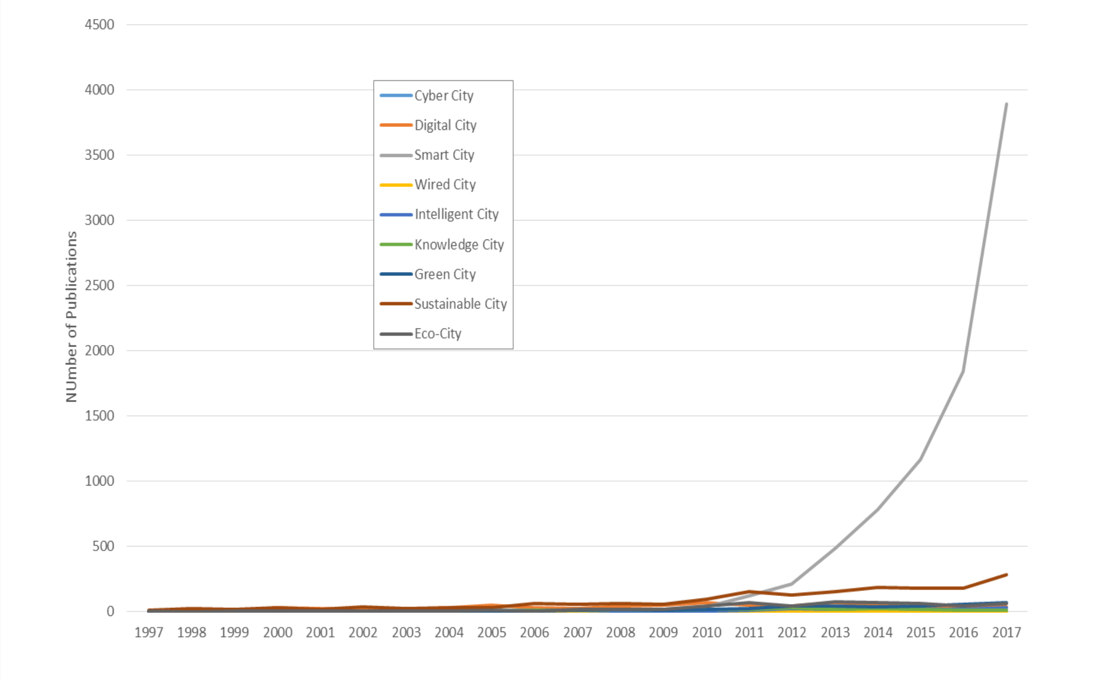


Figure 2-6 Results of Scopus search for TITLE-AB-KEY ("Smart City") and equivalent

2.2.1.1 *Smart city initiatives*

In April 2016, Amsterdam has been elected Europe's Capital of Innovation by the European Commission. This is the result of great efforts made by the city since 2009 to become data-driven [103]. In his review, Fitzgerald has identified the driving forces that initiated such model [104]: (a) the support of group of activities, projects and partnerships such as the Amsterdam's Smart City (AMC) Initiative, a public-private structure that gathers projects on smart mobility, living, society, etc, under a single platform. Other innovation hubs include the Amsterdam Institute for Advanced Metropolitan Solutions, where academic research institutions work along municipalities and companies to run pilots, coordinate efforts and educate students. (b) The creation of city manager posts that are in charge of technologies and smart development. In 2014, the city of Amsterdam has opened the position of Amsterdam's chief technology officer to comply with their will to integrate cutting-edge technologies for a smart city. The objective is to have an official that seeks partnerships and coordinates efforts of the different actors that share the Amsterdam Smart City vision. (c) The multiplicity of projects is also a key aspect. The AMC totals more than a hundred projects dealing with many domains of the urban system since its creation, 9 years ago. This makes Amsterdam a centre for innovation and experimentation. (d) In many projects, citizens have an active role to play in the collection of data, the implementation of new strategies or for advisory purposes. Many examples have been presented where citizens proceeded to waste sorting, gave feedback via social media or simply interacted with the AMC online platform.

Since 2015, Barcelona has often been cited as one of the world smartest cities [105]–[107]. The city first initiated such technological approach in 1992, when it hosted the Olympics. During that period, great efforts have been made to bring the city at the forefront of innovation with, for instance, the implementation of fibre-optic communication [108]. Since then, the city has been a hub for innovation and is nowadays a leading actor in the smart city paradigm. Barcelona's strategy is clustered into four main categories : "Smart Governance", "Smart Economy", "Smart Living" and "Smart People" [109]. Barcelona therefore encourages public-private collaborations on such topics, bringing together multi-national companies, start-ups, academic and public institutions [110]. In 2015, the public-private initiative toward the smart city as led to the creation of 22 programs on telecommunication, smart lighting, smart mobility, renaturation, open government etc, encompassing around 200 projects [111]. Projects such as new grid bus system, Bicing city bikes, smart waste management, smart car parks and sustainable energy systems are showcasing Barcelona Smart City initiative [112]. Finally, the political inclination of the city of Barcelona toward a societal popular movement has triggered a democratic approach to smart city development by encouraging new forms of governmental cooperation by putting the citizen at the core of the decision-making process and by giving to citizens sovereignty over the data produced [113].

Singapore is another city often cited at the top smart cities worldwide [105]–[107]. In 2014, the Singapore government has initiated the Smart Nation initiative [114]. The nation has given a great importance to digital technologies in the past and has now the advantage to lean on a strong digital readiness with a mobile phone penetration of around 150%, 87% of households having a computer and 91% of them having an access to internet via optical fibre, being ranked 1st in term of 4G speed [115], [116]. Moreover, the digital readiness plan goes beyond the simple digital access by considering digital literacy and providing people with the required skills, by promoting digital participation to empower the community and by including everyone via simple design understandable by all [115]. In 2018, the Singapore government has issued its digital government blueprint report that presents its strategies and objectives [117]. Citizens, businesses and public officers are identified as the three main actors participating in the system which aims to be easy to use, seamless, secure, reliable, and relevant, with digitally enabled workplaces and confident workforces. In this prospect, six levering strategies have been initiated: (1) user-centred services, (2) incorporation of policies, operations and technologies in a holistic approach, (3) construction of interoperable systems and platforms, (4) technologies ensuring security and resilience to cyber threats, (5) training and provision of a skilled workforce on digital contents, and, (6) collaborative approach for the co-creation of solutions. However, if Singapore has reached the top of the world's smartest cities, it does not go without criticisms. Indeed, unlike Barcelona that has put privacy and citizen protection at the core of their approach, Singapore has a more opaque model. In its quest for sustainability and quality of life, Singapore proceeds with intrusive systems with little regard for privacy. Examples are in-home movement tracker [118] or surveillance with facial recognition algorithm [119]. This adds up to rather lax personal data protection laws and a tendency for Singaporeans to accept such system [120].

2.2.1.2 *Future perspectives*

The Smart City market is estimated at USD 424.68 Billion in 2017 [121]. By 2025, it is expected to reach USD 2.57 trillion according to a new report by Grand View Research, Inc. with a CAGR of around 18.4% over the 2018-2025 period [122]. In 2016, there were around 16 billion connected devices worldwide for which fixed and mobile phones, computer and tablets accounted for around 60% [123]. A shift has taken place in 2018 where, for the first time, the number of IoT devices will equal the number of "conventional" devices with 10 billion devices out of 20 billion[123]. By 2022, the number of connected devices will reach 29 billion with IoT that will clearly outnumber mobile phones, computers and tablets [123]. The Asia-Pacific region will remain the fastest growing region in the field of smart energies. Additionally, China will account for half of the smart cities in the region by 2025, harvesting around USD 320 billion for China's economy [124]. Latin America biggest cities such as Bogota, Mexico or Buenos Aires will initiate smart city development and Brazil will become the largest investor with projects counting for USD 3.2 billion by 2021 [124]. North America quickly catch up the smart city movement and its smart building market will be estimated at USD 5.74 billion by 2020 [124]. Finally, Europe will encompass the greatest number of smart city initiatives supported by the European Commission incentives for such systems.

2.2.2 Urban sustainability assessment: the importance of data in the smart city paradigm

The accurate assessment of urban sustainability requires highly effective technologies that are able to reflect the actual state of the world in real or near real-time. Several cutting edge technologies are believed to trigger a paradigm shift in catching complex insights of a city organism and of the people that live within. Those technologies can be classified into several categories following the extent to which they actively interact with the surrounding environment. The first category relates to the sensing of phenomena via the use of various sensing and metering devices. The growing interest in and implementation of those devices is leading to the creation of a vast network of sensors known as the Internet of Things. The second is formed of GIS and BIM models and how open data models coupled with accurate satellite imagery can activate an effective collaboration for the creation of up-to-date real-world information. Finally, the last category is defined by software technologies that enable the mining of insightful information within the massive data mine that is the Internet. It implies the design of methods to gather, cluster, analyse and distribute data effectively. Therefore, the concept of crowdsourcing and open data on the Internet is equally important in such a line.

2.2.2.1 *Internet of Things: the sensing world*

A strong emphasis is then given to information as a support to deal with real life urban challenges such as environmental sustainability, socioeconomic innovation, governance, better public services, planning and collaborative decision-making [125]. ICTs are increasingly integrated into everyday life and applications are numerous, including smart homes [126], grids [127], transportation [128], healthcare [129] and cities [130], [131]. The performance of such a model not only relies on the quality of the sensor and meters but also on the reliability and efficiency of information and knowledge exchange[131]. This global network of interconnected objects is called the Internet of Things (IoT). The IoT is described as:

> *"a system that deals with the interconnection of "Things". The word "Thing" refers to any physical object that is relevant from a user or application perspective"*[30].

Some important features are the connection of things to the internet, uniquely identifiable things, ubiquity, sensing/actuation capability, embedded intelligence, interoperable communication capability, self-configurability, programmability [30]. Figure 2-7 briefly represents how the IoT operates at the urban level in order to raise "intelligence", to create a smart city.

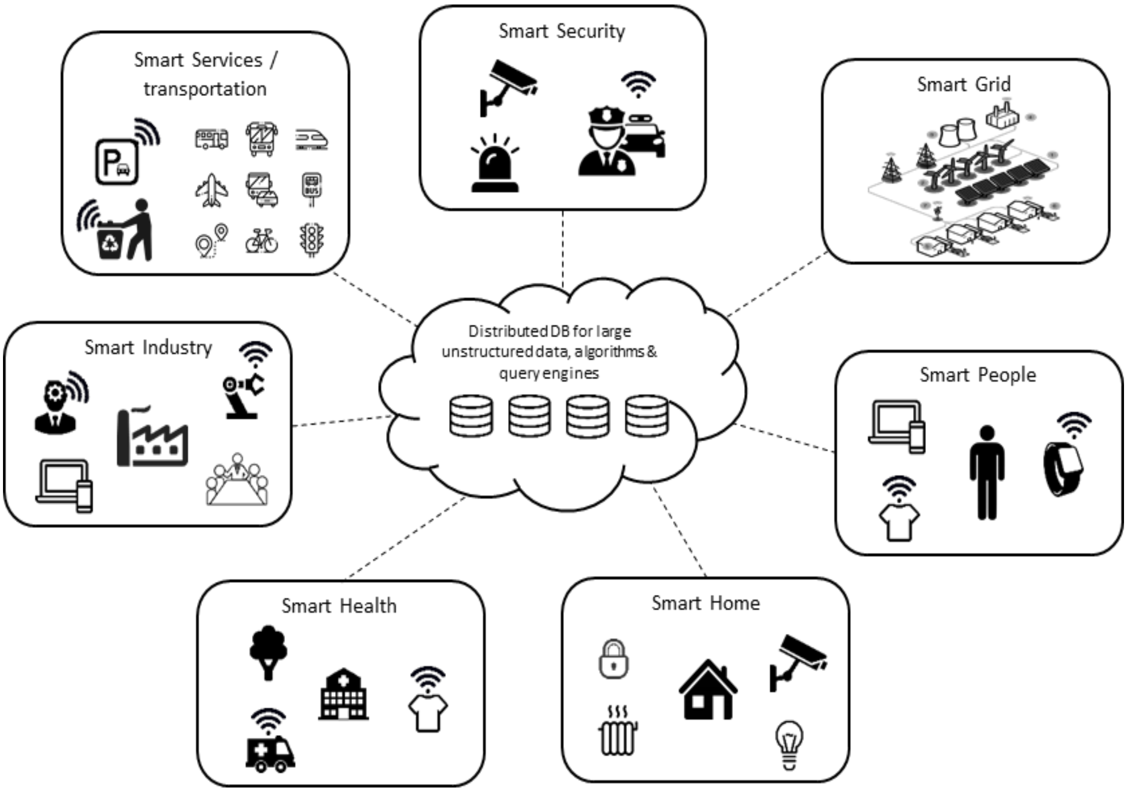


Figure 2-7 IoT representation for urban analytics

The IoT is composed of 3 main phases namely collection phase, transmission phase and processing, managing and utilisation phase [132].

❖ *Collection phase*

The collection phase refers to the sensing and capture of physical phenomenon and the delivery of information via hardware devices such as sensors and communication devices.

This section will mainly refer to the "Things" of the "Internet of Things" that are the sensors. There are various sensors that monitor sensitive phenomena which give significant insights on the city. The "things" can be familiar objects that you own in your home, embedded in factory equipment or part of the fabric of the city. They can be new products and devices built for this specific purpose.

-*Smart homes* include smart devices enabling the measurement of occupancy, indoor environment, behaviours, activities, HVAC efficiency etc [133]. Currently intelligent assistants are the most commonly encountered devices for smart home but the market has the potential to grow fast with applications such as smart control over connected appliances like lighting, heating and air conditioning or white goods; smart security with CCTV and sensors for movement detection, fire and heat detection sensors, gas or water leaks; smart entertainment and information with interconnected multimedia platforms etc [134].

-*Smart transportation* can measure the traffic flow in real-time allowing better traffic management. Traffic management is the biggest challenge of smart transportation with the inclusion of numerous modes of transport and intensive information. This includes tracking public transports and private motorised and non-motorised vehicles as well as the stock of available parking slots, bicycle racks etc. Additionally, safety is another major concern and road conditions, vehicle speed or traffic incidents are some of the aspects that are part of the smart transportation domain. The technologies employed mainly consist of CCTV, radar, LiDAR, GPS. To give some example, efforts made by Tahmid and Hossain for real-time traffic control from digital images [135]; smart parking slots with ultrasonic sensors that detect availability [136]; the study of Aubry et al. in developing a mobile application to report traffic offences [137].

- the objectives of *Smart healthcare* are twofold : preventing possible health issues for anticipated actions and providing better health services by avoiding pointless hospitalisations [138]. Smart healthcare services favours a better understanding of health problems and consequently improved personalisation of treatments. Moreover, a great incentive for the implementation of such systems is cost reduction, a highly appreciated benefit in a domain where costs are substantial. Smart health technologies include wireless body area networks (WBANs) along with mobile apps, which enable to monitor an individual's vital signs or health [139]. Furthermore, some smart home technologies fit into the smart health domain as Lemlouma et al. have demonstrated with the use of motion sensors, smart water and energy meters and/or smart appliance for the evaluation of elderly autonomy or dependency[140]. Another example can include the use of remote sensing technologies for epidemiological studies [141].

-*Smart infrastructures* where water consumption, quality and leakage are monitored along with the possibility to measure relevant events on the electricity network (smart grid), electricity and power consumption, peak load etc [133]. Additionally, smart infrastructures cover the detection of construction structural issues, materials flow and infrastructure maintenance related problems. Even though the publications on smart grid are predominant [127], [142]–[147], other applications exist, such as smartPipes that detect leaks with pressure sensing devices [148] or sets of sensors to detect bridges structural deformation for Infrastructure Health Management [149].

-*Smart services* monitor citizens' experience and satisfaction with public services in order to deliver the best-quality services possible. Detection of natural hazards or crimes, tracking waste volume for collection or counting facilities occupancy are few examples of the aspects covered by smart services.

❖ *Transmission phase*

The transmission phase comprises means to deliver the collected information to application, services and external servers. It refers to the other notion of the "internet of things" which is the "internet"; a network connecting all the different devices one with another and allowing a simple and fast communication of information and knowledge. This network is made possible by the implementation of cutting edge communication technologies increasingly efficient and reliable.

- *Radio-frequency identification (RFID)* is based on electromagnetic fields to automatically identify and track tags attached to objects. The RFID technology can be used to identify virtually any object, including animals, clothes, and even human beings [131]. It has already been implemented to track food products in supply chains [150], construction materials and building components in construction projects [151], waste for real-time waste management [152] or animals displacements [153].

-*Wireless sensor network (WSN)* is a network of small, low-cost and low-power devices that are connected one to another. This network has the advantage of being economical and simple and provide a good opportunity for devices connection [131]. The devices can measure a variety of environmental and physical properties as well as play the role of actuator for control. The advantage of a wireless system is the opportunity to deploy it on a large scale in a non-intrusive and cost effective way, an essential task in the smart city paradigm.

- *Wireless Local Area Network (WLAN) such as WiFi, Ultra-wideband, ZigBee, and Bluetooth* are certainly the most widely used short-range wireless communication technologies allowing access to the internet or communication between devices. They play a key role in data transmission and communication between sensors.

-*4G (LTE), LTE-A, and 5G* are standard high-speed wireless communication based on GSM/EDGE network technologies. 4G and LTE-A (also called 4G+) are currently the most widely spread standards reaching a data transmission speed of 100 to 200 Mbps. The 5th generation of mobile network is planned for 2020 and will supports bandwidth of up to 10

Gbps [154]. This last generation has the advantage to be designed to support the IoT by incorporating machine-to-machine (M2M) communication.

- *Processing, Managing And Utilisation Phase*

Once all the devices are connected, they will certainly produce a large amount of data. The integration of the IoT significantly increase the amount of data that need to be processed and stored [125], [155], [156]. Therefore, Cloud computing is a good option in order to deal with such challenge [125], [157]. Mell & Grance define cloud computing as:

> *"a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction*" [131], [158].

Cloud computing provides services such as platform as a service (PaaS), software as a service (SaaS), and infrastructure as a service (IaaS). Cloud computing has been greatly used by experts for the deployment of services [139], [159]–[161].

An alternative to cloud computing increasingly considered is the fog computing paradigm [162]. Fog computing extends cloud computing capacity/capability by integrating the edge of the network. It includes wide, dense, distributed edge devices and datacentres that can serve application that requires low latency. Additionally, such approach enables to differentiate data and services sent to the edge and cloud networks, lowering the traffic to the cloud. In such perspective, requests sent by IoT devices can be served by the edge of the network when required rather than in the cloud. Finally, fog computing addresses scalability challenges with an increasing amount of endpoint, reducing processing on the cloud.

Beyond storage and processing capacities, this phase includes mechanisms for abstracting pieces of information, automatically and dynamically discover them and aggregate services from "basic" once [132].

2.2.2.2 *BIM and GIS technologies*

BIM and GIS models and databases are valued knowledge resources that, via an exhaustive and detail indexation of building to urban level objects, can help in the determination of certain key performance indicator for sustainability assessment. This section present those assets and a sustainability assessment can leverage on them.

❖ *BIM technologies*

BIM has been originally used by the AEC industry as a design tool allowing a 3D visualisation of the different building components [163]. It is for instance used for the coordination of Mechanical, electrical and plumbing (MEP) elements to ensure there are no clashes [164]. If BIM has gained a great interest throughout the years for building design, its implementation on existing structures remains limited [165]. The BIM field has seen some enhancement of its information embedding, leading to more practical uses. The information is becoming richer and a new concept has been introduced, the "nD BIM". The "nD BIM" goes beyond the simple 3D BIM with the addition of new dimensions as shown in Figure 2-8.

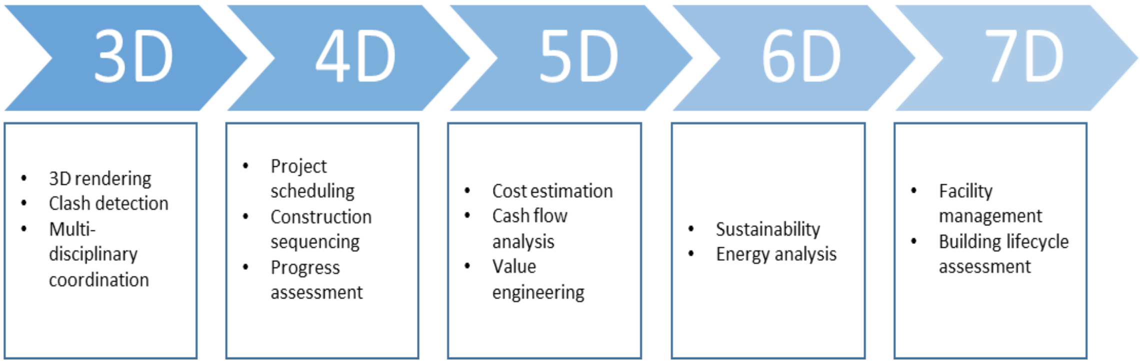


Figure 2-8 the BIM dimensions (nD BIM)

In the context of sustainability assessment, the 5th and 6th dimensions are particularly interesting since they validate the use of BIM models for estimating environmental and economic metrics. Some research concerning the use of 5D BIM has been pursued: Lu et al. proposed a 5D BIM framework for cash flow analysis and project financing along with the possibility to run what-if scenarios for alternatives analysis [166]; Cheung et al. developed an automated cost estimation model based on the building geometry [167]; Smith draws guidelines for a successful implementation of 5D BIM by the industry after spotting the issues that limit its adoption [168]. The usage of 6D BIM has been demonstrated by Yung and Wang who have designed a model performing building sustainability assessment automatically [169]; Jalaei and Jrade use BIM model to evaluate and allowed credits in compliance with the LEED-NC sustainability rating scheme [170].

An alternative to the Nd BIM classification currently used by the British government to develop its BIM strategy is the 3 maturity levels of BIM [171]. The BIM maturity model starts with the level 0 with simple CAD, drawings and documents. BIM level 1 includes 2D and 3D BIM models and sets standards for digital indexation of construction information. BIM level 2 describes a level of maturity with the inclusion of collaborative and federated models, the creation of a common data environment for construction related information. Finally, BIM level 3 is yet to be defined but is thought as an online, collaborative and inter-disciplinary information models for the AEC that enables a seamless information exchange between heterogeneous services and organisations. The British government has for target to reach BIM level 2 by 2016 and its consolidation and full adoption by 2020, and is currently developing an agenda to reach BIM level 3 in a near future [172].

Currently, the BIM has to overcome some technical challenges to trigger its full potential. Indeed, Volk et al. have identified 3 major challenges that need focus: (a) the automated creation of BIM model along with the capture of supporting data; (b) the update and maintenance of the BIM models; (c) the representation of uncertain objects and relations from already existing building into BIM [165]. Additionally, BIM has shortcomings in interoperability and heterogeneous applications, issue covered in section 2.2.3.2 [173], [174]. Therefore, BIM models must overcome those issues, in particular, real-time update and interoperability, in order to estimate accurate urban sustainability KPIs.

The shift to a real-time automated as-is BIM model is believed to enlarge its use from design to the construction and operative stages; to become a life-cycle asset [164]. At the construction stage, real-time as-is BIM models would support construction planning and operations, tracking material and technical needs [163]. At the operative stage, real-time as-is BIM would allow the monitoring of the building structure for health and safety measures as well as the management and maintenance of MEP components [163]. The real-time synchronisation of building information is enabled by the integration of technologies that collect field data automatically or semi-automatically. Many efforts have been done in the collection and automatization of as-is BIM model such as the incorporation of Radio Frequency Identification (RFID) and barcoding during the construction process that tracks building components and appliances [151], [175], [176] or 3D scanning (often Light Detection and Ranging LiDAR or Laser Detection And Ranging LADAR technologies) of existing building that capture building geometries and equipment [177]–[180]. Other technologies include high resolution camera and images [181]–[183], and photogrammetry [184], [185].

- *GIS technologies*

During the past two decades, GIS skyrocketed both technically and in its use. Those improvements have been triggered by the increasing computation power and by the intensive use of new technologies such as global position systems (GPS) and aerial and satellite imagery [186]. Currently, studies on GIS and spatial analysis are often associated with the use of remote sensing technologies. Remote sensing goes beyond the simple pictorial capture of the land by enabling the capture of a bigger portion of the electromagnetic spectrum. This allows the acquisition of insightful information concerning the land observed [187]. To cite a few: Hegazy and Kaloop have studied the evaluation of Daqahlia governorate urban sprawl comparing successive satellite remote sensing data [188]; Contreras et al. have focused their research on monitoring building recovery after disasters (an earthquake) using remote sensing [189]; Gandhi et al. have used remote sensing and GIS for the evaluation of the vegetation changes over time [190]. Additionally, satellite imagery has reached a resolution of about 0.31 meters that allows the visualisation and identification of smaller and smaller objects [191]. Indeed, high-resolution images along with the advance in machine learning algorithms have achieved significant improvement in object detection and classification [192]. For instance, Bai et al. have successfully used a support vector machine (SVM) for the detection of airplanes, ships, houses, stadiums, bridges, and vegetation [193]; Li et al. have achieved a robust rooftop extraction with a conditional random field (CRF) [194]; Wang et al. have developed a convolutional neural networks to accurately capture road networks [195]. Other remote sensing technologies include Light Detection And Ranging (LiDAR) observation from the ground or sky. Those technologies are increasingly used within the GIS community for the capture of topographies and objects. Many uses can be found in the literature such as Gagnon et al that determine rooftop PV potential from LiDAR data [196]; Mitasova et al that assess dune migration and coastal erosion with a topographic analysis based on LiDAR [197]; or Shirowzhan and Trinder that utilises LiDAR and SVM technologies to classify building in 3D urban models [198].

LiDAR or photogrammetry are cutting-edge technologies that have triggered the development of a new type of GIS, the 3D GIS [199]. 3D GIS introduced the possibility for 3D visualisation and three-dimensional spatial analysis of the cityscape. Several schemas have been created to store 3D attributes of urban objects such a COLLADA [200], Keyhole Mark-up Language (KML) [201] or the Geography Mark-up Language (GML) [202]. The CityGML data model, an application schema for GML, is often cited [203]–[209] to support

3D GIS. It has been developed by SIG3D (Special Interest Group 3D) and adopted by the Open Geospatial Consortium (OGC) as a formal data standard in 2008 [205]. This semantic model is composed of several, gradually fine, levels of details (LoD) of the urban objects. The LoDs express different degrees of abstraction of real-world objects (often buildings) and with its different potential applications of 3D GIS [208].

3D GIS has proven to be useful for various applications: Bremer et al. have used 3D GIS for the modelling of solar irradiation to support solar panels integration [206]; Campanaro et al. worked with 3D GIS technologies to study architectural structures for cultural heritage preservation management [199]; Tymkow et al. simulated flood phenomena based on 3D spatial information [203]; Landeschi et al. have used 3D GIS for visual analysis and artefact pattern detection [210]. Additionally, 3D GIS can bridge two critical domains that are GIS and BIM technologies proving its value for interoperability. This last feature will be described in more details in Section 2.3.3.3.

Equally, 4D GIS models are being investigated. 4D GIS is an extension of the 3D GIS with the inclusion of temporal data [211]. In this approach, data acquisition is done in different points in time which allows the analysis of territorial and structural changes [212]. Geometric data acquisition and map update are important challenges in the creation of accurate 4D GIS models. Daily 3D scans and high-resolution images have been considered for such task but still present some concerns especially on the storage of subsequent massive 3D datasets [213]. Real-time kinematic Lidar (RTK-GPS) have been considered as well to track changes at the sub-centimetre to centimetre levels and are promising technologies for a fine representation of the model variations [213], [214]. In the literature, 4D GIS are implemented for project management and construction monitoring [215], [216] as well as for facility management [211] or environmental analysis [212].

Finally, the implementation of free and open GIS resources are believed to democratize GIS technologies and to capitalize on public participation and users collaboration [217]. The Open Geospatial Consortium (OGC) is the reference organisation for the development of open standards for Geospatial data [218]. Those standards include the previously cited Geography Mark-up Language (GML), CityGML and Keyhole Mark-up Language. OpenGIS along with web map servers that provide web services is a trending approach that enables the acquisition, processing and sharing of spatio-temporal information within a unified system [219]. Hu et al. have used OpenGIS and web map server for the deployment of a geospatial web service for agriculture soil moisture, interoperable with other agricultural applications [220]; Prietro et al. have produced 3 web decision support tool for energy retrofitting solutions based on OpenGIS standards such as CityGML [221]; Schröter et al. proceed to a flood risk assessment of the city of Dresden, Germany using already existing CityGML model of the city [222].

2.2.2.3 *Data mining and crowdsourcing*

- *Data mining*

PhridviRaj and GuruRao define data mining as [223]:

> "*A process of discovering hidden patterns and information from the existing data*."

It is not a recent concept and the approach of "fishing through data" has already been greatly studied in the past decades. However, modern development in data analytical models, machine learning and database technologies along with the exponential increase of data volume (Big Data) have led the field to gain in popularity [224]. In practice, data mining is used to identify and classify non-trivial patterns for the retrieval of insightful information and/or the creation of models from a vast unidentifiable mass of data. Following the mining technique employed, various structured, unstructured or semi-structured data types can be handled such as text, tabular data, image, video, audio, hypertext etc [225]. Figure 2-9 shows the various techniques employed within the data mining field (aka data science) for exploration and modelling [226]. Many domains are covered such a statistic with the use of Chi-square test, ANOVA, variance, covariance, regressions etc; machine learning with decision trees, artificial neural networks, support vector machines, Naive Bayesian classifiers, k-nearest neighbours or Self-organizing Map etc; and other including hierarchical clustering, k-means clustering etc.

Moreover, there is currently a tremendous amount of data produced from heterogeneous sources each day. Sources include [227]:

- *World Wide Web*: There are billions of online documents and web pages that can be mined as well as social media feed. Additionally, the "hidden" part of the internet can be mined which contains servers logs or the Web graph itself.
- *The Internet of Things*: The rise of sensor networks has enabled the intensive release of data that represent a valuable source of information for mining. From smartphone to CCTV, sources are varied and can cover a large span of domains.

- *Financial market*: Data mining is already actively used to analyse and predict stock, bond, money markets etc. On a smaller scale, everyday life transaction such as withdrawal, credit card activity, money transfers can be mined for the discovery of fraud or uncommon activities.
- *User input*: All type of user inputs that do not require the use of the World Wide Web. For example, telecommunication records that companies often analyse for maintenance and marketing purposes.

In the context of urban sustainability assessment, data mining can effectively support the acquisition of information on user experience with public services, market changes, demographics, emergency calls or organised events. For instance, Güllüoğlu has shown how data mining techniques could help in the customers' segmentation of a Turkish supermarket chain, showcasing their use for demographics [228]. Both Xiang et al. and Wang et al. have demonstrated the use of data mining to capture the user experience and satisfaction of hotels and clothing shops respectively via online text mining [229], [230].

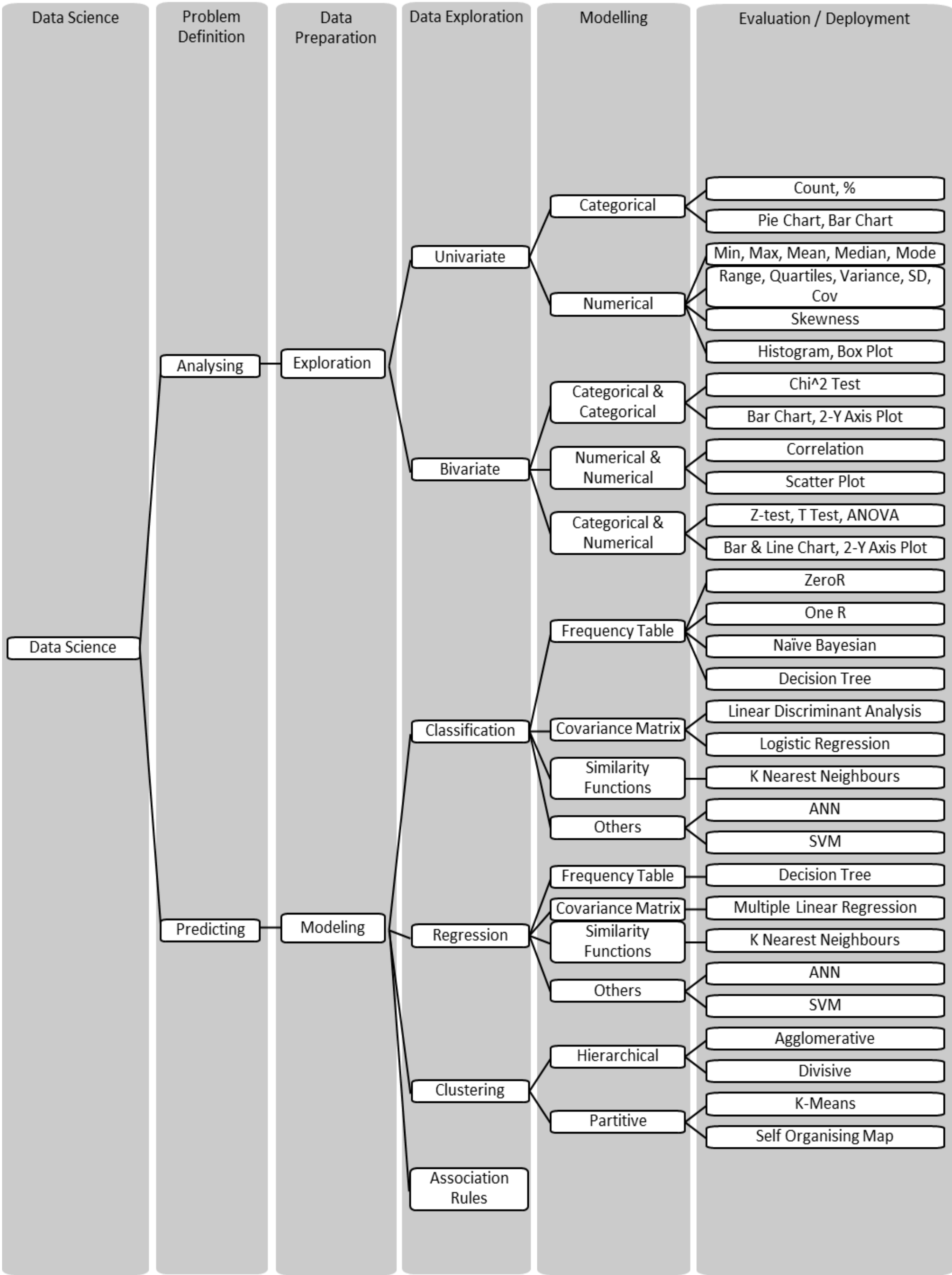


Figure 2-9 Data mining map based on [226].

❖ *Crowdsourcing*

User inputs and World Wide Web mining have introduced a new notion of knowledge acquisition from the "crowd". This notion has been studied and a new approach has emerged in the past decade known as "crowdsourcing". As Howe put it in his 2006 Wired article [231]:

> *"Crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. This can take the form of peer-production (when the job is performed collaboratively), but is also often undertaken by sole individuals. The crucial prerequisite is the use of the open call format and the large network of potential labourers."*

This early definition has since been extended thanks to the rise of social networks and platforms. Indeed, additionally to the original conception of crowdsourcing that can be labelled as "active" crowdsourcing, "passive" crowdsourcing has emerged based on information provided on social media platforms, discussion forums and smartphone apps [232]. The main difference between active and passive crowdsourcing lies in the incentive to participate in the task. Active crowdsourcing is directly called to perform discrete tasks and is sometimes rewarded via platforms such as Amazon Mechanical Turk or Figure Eight.

The benefit of crowdsourcing is twofold:

- *The Wisdom of crowds:* It is believed that the multiplicity of sources associated with heterogeneous profiles have the potential to lessen biases and increase accuracy [232]. This notion has been framed in great extent by Surowiecki in his book "The Wisdom of Crowds" where he showcases problems solving and decisions making in various field such as biology, computer science or politics [233] (with all due respect to Charles Mackay and Gustave Le Bon that argued differently in the XIXth century [234], [235]).
- *Focus:* Individuals in the crowd are performing specific tasks which allow the production of a high-quality knowledge base. A better focus in the dataset will, therefore, lead to less time-consuming data mining and reduce costs [236].

Crowdsourcing along with data mining technologies have been studied and applied for real cases: Domdouzis et al. present the ATHENA system, a crisis management system that integrates social media features (mobile app, news feed, geolocalisation) to prevent criminal activities during crisis situation [237]; Erraguntla et al. have developed the MySidewalk™ mobile application that asks its users to pin sidewalk condition issues for inventory [238]; Nik-Bakht and El-Diraby have implemented crowdsourcing in form of a Game With A Purpose to draw an urban sustainability framework from the public perspective, an interesting initiative that put the people into the creation loop [239].

2.2.3 Technological Issues related to the Smart city

The broad review of the literature has drawn an interesting perspective on how the authors refer the issues related to Big Data and Internet of Things as "challenges" [18], [125], [131], [155], [240]–[243]. The term "challenge" involves the fact that those issues must be overcome. It prefigures that this paradigm is already widely accepted within the expert community and that its implementation is certain. Thus, leading researches in the domain are toward the development of solutions to overcome those challenges. The issues are multiple and spread across various disciplines. This section will cover technical issues and solutions addressed in the literature concerning IoT and Big Data for Smart Cities.

2.2.3.1 *The IoT Technical issues*

The field of Big Data is growing as exponentially fast as ICTs. The final goal of big data, which is the reliable transmission of a large amount of information instantly, requires the development of increasingly powerful hardware and software. It is one of the major challenges that the current research area faces. The technological limits do not allow the current system to fulfil expectations [131].

- *Network Architecture*
  The IoT must overcome certain challenges related to its architecture. Indeed, the network of interconnected devices must be flexible with the adoption of a plug'n'play approach. The architecture must ease the integration of new nodes as well as the upgrade of already existing ones. It implies perfect interoperability of the systems for a seamless communication between the network nodes [21].
- *Energy efficiency*
  Wireless moving sensors are not linked to any power supply and must be self-sufficient. This requires the sensors to be energy efficient for long-lasting discontinuous information transmission [20]. Additionally, in an environmentally sensitive context, IoT is valuable only if it is constrained by environmental requirements which make energy efficiency an imperative.
- *Privacy and Security*

Data holding sensitive information must be protected from eventual hacks. National security, enterprise secrets and individual privacy must be preserved [244]. Thus, efficient services, applied on a large scale, should ensure resilience to attacks, data authentication and client privacy [245].

– *Storage and Processing Issues*
  With data continuously produced from various sources, the current database technologies are insufficient to deal with the amount of data at a quick processing rate [131]. An improvement track is the use of the cloud to deal with storage limits. However, uploading this large amount of data will take a large amount of time which is not compatible with fast-changing data [240].
– *Data integrity*
  In order to ensure the good use of Big Data, the system must be: fault-tolerant, allowing an "acceptable" level of failure; scalable in order to meet a substantial workload; flexible to answer to different queries and operations from heterogeneous data format; and reliable.
– *Quality of Service*
  Big Data in smart cities being based on a variety of protocols and technologies and a heterogeneous network, providing the right QoS for the entire network is a real challenge [246]. Services must respond to the various application requirement without compromising reliability, flexibility, scalability and fault-tolerance of the network [131].

2.2.3.2 *Interoperability issues*

Interoperability between heterogeneous information systems is equally important for information transmission and the provision of increasingly intelligent services. The key is the creation of a ubiquitous data landscape where the information is unbounded by formats and structures.

❖ *Interoperable IoT*

The IoT is growing fast with the inclusion of more and more sensors and information. It supports many different applications, as shown in Section 2.2.2.1: to name a few smart wearables, transports, surveillance, health care, industry, energy etc.

Data heterogeneity is a barrier to the creation of a ubiquitous IoT eco-system [29]. In other words, sensors data interoperability is essential to fully take advantage of the data produced by the IoT. However, with the IoT application skyrocketing and covering multiple independent and uncoordinated domains and organisations, the creation of a traditional standard data model is an impossible task [247]. It is most likely that predominant data models will remain in their respective fields of use. Therefore, efforts on interoperability should not focus on the development of common data and meta-data formats but rather to ensure machine translation or understanding of formats across systems [247]. Ultimately, this will ensure Machine-to-Machine (M2M) communications, an essential feature of the IoT vision [132]. From a market perspective, McKinsey has estimated that IoT will have $11.1 trillion per year economic impact by 2025 and that achieving complete interoperability (not restricted to data but also to protocols and systems structure) would unlock 40% up to 60% potential economic value [248]. Most certainly because a non-interoperable system limits sensor data aggregation for subsequent processing, M2M communication and is time and resources consuming.

❖ *Interoperable BIM*

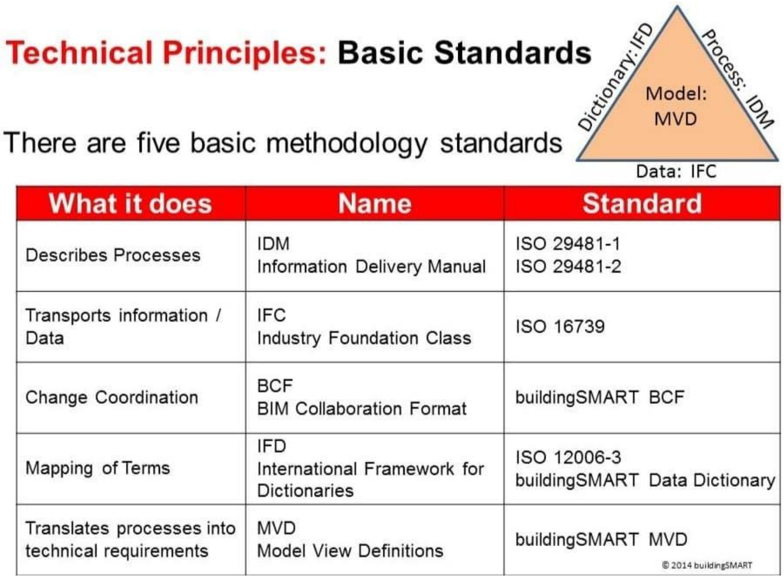
The AEC (architecture, engineering, construction) industry is a split sector where several actors from different domains are required to cooperate on a project within the time and budget allocated [249]. The industry is therefore information-based and knowledge-intensive [249], [250] and communication between different parties is key to a successful realisation. It is essential to provide the good means for an effective communication. In 2018, PlanGrid partnered with FMI to survey the costs of poor data and miscommunication within the AEC industry [251]. They estimated that poor data and miscommunication accounted for about $177.5 billion in the U.S in 2018 and 48% of all rework. Respondents stated that the top causes for losses are erroneous or incorrect project data for 34.4% of them and the difficulty accessing needed project data for 23.8%.

From sophisticated document management systems that assess construction progress to personnel video tracking on construction sites [249], [252], the AEC industry has seen many collaborative technologies for an effective integration of actors, processes, business systems and information throughout the entire project life cycle [253]. BIM has acquired a central role in the construction supply chain by providing schedules that allow a good asset and process management (people, materials, transportation, storage etc) [254], [255]. In a context where the AEC industry has a really disparate knowledge across domains, BIM enables the storage and management of information acting like a database where humans or machines can extract knowledge [249]. BIM along with the increasing integration of information and communication technologies (ICT) are believed to trigger a new paradigm of knowledge management in AEC where semantic gains are importance [249], [256].

Unfortunately, specialists estimate that BIM is not mature enough for complete interoperability among heterogeneous applications [173], [174]. Consequently, studies have shown that interoperability issues have a cost (around $15 billion per year in the U.S. in 2004 [257]) and limit the BIM adoption by the industry [258]. As Muller exposed in his article[259], there are multiple concerns associated with interoperability:

- *Business concerns:* the adoption of BIM by a company is by nature a strategic action which requires all the stakeholder to be involved
- *Process concerns:* the choice of BIM redefines and centralizes the processes requirements for construction, design, and operation
- *Services concerns:* all the services from various companies must work together around a single knowledge base
- *Data concerns*: the ability for machines and systems to work together around a common digital resource.

BIM for urban metrics determination would benefit from cloud-based technologies and open format. Cloud-based BIM model would ease the reach, manipulation and update of models without any need for commercial software interface [164]. Moreover, it has great potential for the elaboration of collaborative solutions in real-time, for the centralization of intensive information and for visualisation [260]. Along with cloud computing, BIM must perfect its interoperability with the use of an open format such as the IFC format. In that context, BuildingSMART has initiated the creation of an openBIM framework that relies on 5 standards as shown in Figure 2-10 [261], [262].



What it does	Name	Standard
Describes Processes	IDM Information Delivery Manual	ISO 29481-1 ISO 29481-2
Transports information / Data	IFC Industry Foundation Class	ISO 16739
Change Coordination	BCF BIM Collaboration Format	buildingSMART BCF
Mapping of Terms	IFD International Framework for Dictionaries	ISO 12006-3 buildingSMART Data Dictionary
Translates processes into technical requirements	MVD Model View Definitions	buildingSMART MVD

Figure 2-10 openBIM standards [262]

Many examples in the literature have already demonstrated the potential of cloud computing and IFC for the creation of “open BIM platforms” facilitating information access and exchange [159], [260], [263].

❖ *Interoperable GIS*

The GIS community has raised concerns about the challenge of interoperability, and consequent efforts have been done to tackle the issue. The angle taken to overcome the issue is the creation and inclusion of common standards that dominate the field [264]. The Open Geospatial Consortium OGC is one of the main contributors in the standardisation of the GIS field providing an exhaustive set of open and free standards to the community. Those standards includes data schema such as SensorML for sensor meta-data descriptions , netCFD for space and time-varying phenomenon gridded data, OpenGIS Standard for the Web Processing Service, WKT, GML, cityGML and more. The adoption of those standards is strengthened by their community-driven development, leveraging on user experiences and collaborative projects with open tools such as PostgreSQL/PostGIS spatial DBMS or GeoServer and MapServer web map server [264].

BIM and GIS systems are seen as complementary by many for the description of building to urban level object with BIM models covering rich geometric and semantic information about a buildings life cycle and GIS covering geospatial modelling [207]. Despite their mutual need of information from each other, the two models are the results of uncoordinated efforts. For example, BIM considers local cartesian coordinate system while

GIS rests on geographic coordinates and map projection [265]. Applications of a BIM-GIS vision could include the possibility to track building materials, urban scale heating demand forecasting, noise modelling or more [205]. Consequently, efforts have been deployed in recent years for the creation of interoperable BIM-GIS systems [205], [207], [254], [265]–[268]. Those efforts focuses on the transformation and/or mapping of BIM attributes such as geometries and coordinates in the standard IFC format with their equivalent GIS attributes in the cityGML standard. Transforming BIM to cityGML is often favoured in the literature as BIM models present a greater level of details than GIS models. In a similar fashion, studies have been conducted for the integration of IoT and BIM where IoT attributes stored in RDBMS databases are link to BIM server attributes via a mapping of standardised schemas [269], [270].

Of course, issues related to IoT and big data are not limited to the technical perspective and many managerial, political and ethical issues come into play in the Smart City paradigm. Those will be further discussed in Section 2.4.

2.3 SEMANTIC WEB TECHNOLOGIES TO UNLOCK SMART CITIES POTENTIAL

One of the greatest challenges mentioned earlier is the creation of an interoperable and ubiquitous information system to support the effective implementation of smart cities. In this section, one of the key technological assets considered to solve this issue, the semantic web technologies, is reviewed in detail. The core principles of semantic web technologies will be exposed along with ontologies engineering principles and the fundamental ontologies believed to support the IoT and sustainability assessment.

2.3.1 Background on the Semantic Web

2.3.1.1 *The future of the World Wide Web*

A major leap in ICTs has been defined by Tim Berners-Lee in its 1989 seminal article "*Information Management: A Proposal*" [271] which describe the principle of hypertext and what is now known as the "World Wide Web" (WWW). Numerous applications have then been produced, demonstrating the increasing interest in the World Wide Web and the popularisation of the Internet, a free and open WAN. From there, knowledge could be not only created and shared but also made publicly available to potentially anyone. The WWW has then evolved, integrating social web and real-time information exchanges, producing more information than ever before. In 2017 solely, the overall volume of data generated since the beginning of the digital era has increased by 1.5 [242] and more data have been created in 2014 and 2015 than the entire previous history of the human race [272] (a trend that increases exponentially).

Figure 2-11 demonstrates, the semantic web is often considered as the next step of the WWW, also labelled Web 3.0. Tim Berners-Lee when creating the WWW, originally pictured it as the semantic web [273].

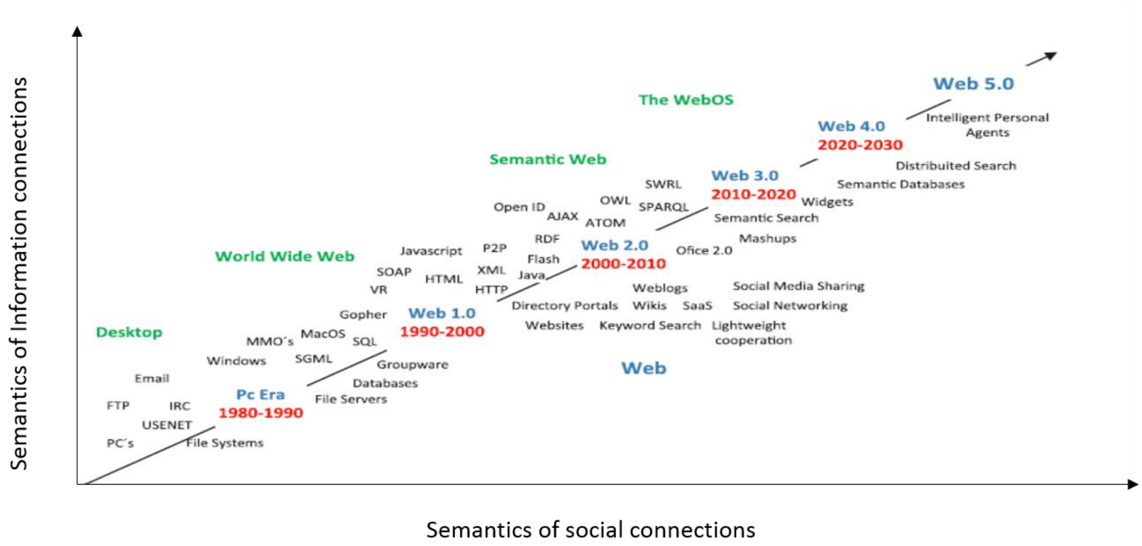


Figure 2-11 World Wide Web evolution [273]

In the current state of the web, information is bounded, encoded into different formats which stop pieces of information to relate one another. The information web must evolve from a space where documents are linked to a space where data are linked [274]. These principles are the core foundation of the semantic web and the linked data paradigm. This will enhance data by the creation of links between pieces of information.

Currently, the most striking example of available linked data resource is the DBpedia Linked dataset [275] which, essentially, transposes the content of Wikipedia into Linked Data. DBPedia is connected to many more resources than Geonames. As of 2011, 295 different datasets were linked to DBpedia. In August 2018, the cloud counted 1,224 identified datasets [275]. This illustrates the participatory approach that is one of the core principles of semantic web. In this new vision, information sources are not duplicated but reused and linked, ultimately creating a space of information where every piece of knowledge is a singularity.

2.3.1.2 *Semantic web vs interoperability*

One of the major benefits of semantic web technologies is that they can produce interoperable systems despite information systems' heterogeneity [276]. Early on the creation of global information systems such as the WWW, interoperability concerns have been issued [277]. Indeed, as the information space got increasingly distributed, autonomous, diverse and dynamic, the semantic and structural heterogeneity of crossdisciplinary, multicultural and multimedia technologies have increased as well [277]. Integration and interoperability have therefore become challenging tasks in the implementation of seamless system communication. Semantic web technologies come into play as they support integration and interoperability by offering extra meaning to the information [278]. Indeed, because any specific domain semantics can be explicitly represented with its various entities, properties and relationships, systems structural heterogeneity does not matter as they are based on a common conceptual representation [276].

To illustrate the usefulness of semantic web technologies for interoperability, the example of collaborative networks within the smart city paradigm has been chosen. Indeed, semantic web technologies appeared to be a valuable mean when it comes to collaboration between heterogeneous systems and organisations. The homogenisation of the information has the potential to unify diverse entities into a global system of information. Consequently, such technologies are currently being considered in the development of collaborative networks (CNs) [279]–[281], an emerging paradigm for goal-oriented heterogeneous organisations' collaborations [282]. As Camarinha-Matos and Afsarmanesh define it, CNs are

> *"constituted by a variety of entities (e.g., organizations and people) that are largely autonomous, geographically distributed, and heterogeneous in terms of their: operating environment, culture, social capital, and goals. Nevertheless, these entities collaborate to better achieve common or compatible goals, and whose interactions are supported by computer network."* [283]

In the upcoming paradigm of smart cities and smart services, collaborative networks are believed to improve workforce performance, decrease environmental impact, disseminate information and knowledge, create fast responding systems, accomplish concurrency in operations and promote innovation [284], [285]. CNs have already been implemented in numerous areas such as manufacturing for adaptive and responsive production [286]; construction for an improved management and coordination of large projects [287]; or energy provision to ensure flexibility and sustainability of the service [288], [289]. The latter is actually a good example on how these three fields that are semantic web technologies, collaborative networks and smart energy provision (part of the smart city paradigm) are complementary and how one can benefit from the others. Certainly, the energy industry is fast changing, moving from a centralized system where one entity ensures the energy generation and provision, to a distributed, bidirectional energy generation and distribution [290]. The end goal is to achieve an energy system that integrates more and more renewable energy sources, a fast demand response, flexibility in order to smooth peak loads, the reliability of the network and the inclusion of consumer within the loop, becoming prosumers [291]. New roles will come into play, differencing energy suppliers with the distribution network and system operators, transmission system operators, prosumers and energy aggregator services [290]. In this vision where different entities come to work together in order to provide a globally efficient service, CNs are much needed and so are semantic web technologies. As CNs theory defines it, semantic web technologies will enable seamless information flow between partners, consolidating inter-enterprise knowledge [286] and overcoming heterogeneous data structures [292]. An example found in the literature is the Knoholem project for “Knowledge-based, holistic, energy management of public buildings” where heterogeneous data are linked via the use of an ontology [293]. Local ontologies such as building ontologies are aligned with the core Knoholem ontology, enabling interoperability between systems. This concrete example showcases how the potential of semantic web technologies to deal with interoperability, an essential feature for collaboration in the future smart city landscape.

2.3.2 Introduction to semantic web technologies

2.3.2.1 *Ontologies backbone*

The semantic web has been formalised by a set of standards developed by the World Wide Web Consortium (W3C), defining data language and schema, query language, vocabulary and referencing possible applications [294]. In those standards one of the core concepts of the semantic web can be found which is the concept of ontology. An ontology is:

> "*a specification of a representational vocabulary for a shared domain of discourse — definitions of classes, relations, functions, and other objects —* "[295]

In the particular context of computer science, the term ontology can be somehow misleading. Indeed, there are several levels of understanding behind that term from conceptual to more concrete interpretations [296]. On a conceptual level, an ontology can be characterised as a set of statements that are known to be true. This set of truthful statements is a representation of real-world phenomena and, as such, one can rigorously describe via an ontology a specific domain. Those statements are constructed around entities, properties and their interrelationships. This conceptual representation can be formalised in a computer readable way, which is what the computer scientist most often refers as ontologies. Consequently, if those computer readable concepts are considered as being universally true, then artificial intelligence can extract additional knowledge by inferring information over the defined rules. Moreover, those concepts can be instantiated with real-world objects, and by reasoning over an ontology, all the relationships defined can be applied to the subsequent instances. Therefore, when the conceptual framework is built via the ontology, there is no need to proceed to a case-by-case description of the instances [295].

Before engaging into the description of ontologies and the standard protocol and languages behind them, it is necessary to introduce existing feature such as URI/URL and XML.

URI stands for Uniform Resource Identifier and is described as

> *"compact string of characters for identifying an abstract or physical resource."*[297]

They are strings of characters without spaces that allow targeting a specific resource on the web. It must not be mistaken with Uniform Resource Locator, or URL. Indeed, an URL is a subclass of URI that is identified by its location rather than other attributes it may have. URLs comply with HTTP (Hypertext Transfer Protocol) and its methods [298]. Note that HTTP refers to set of communication transactions (typically GET, POST, DELETE) that can be performed between a client and a server. URLs are composed of a scheme (e.g. HTTP), an authority which labels the server where a resource is sought and a path which locate the resource within the server.

Finally, XML stands for eXtensible Mark-up Language and is an extension of the Standard Generalized Mark-up Language (SGML). They have been designed in order to mark-up information with their meta-data in a computer-readable manner [299].

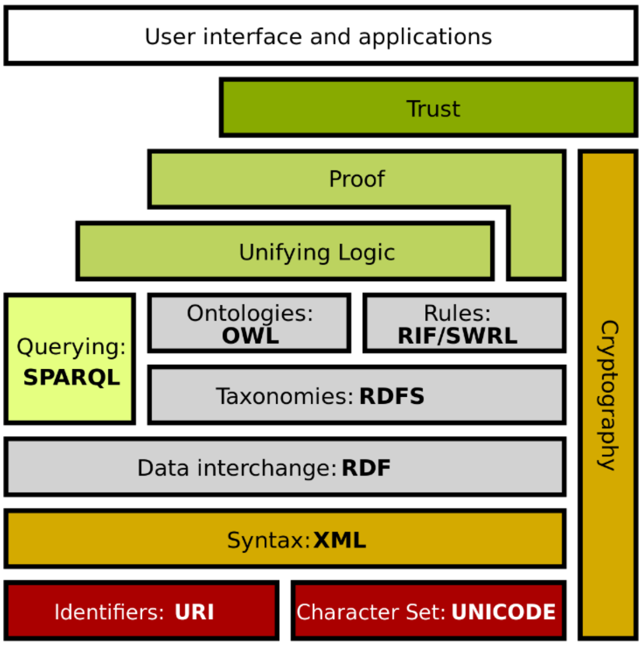


Figure 2-12 Semantic Web Language stack

As Figure 2-12 suggests, RDF, that stands for Resource Description Framework, is defined on top of XML and URI concepts. RDF aims at enriching information about resources on the WWW by adding machine-readable and understandable meta-information [300]. Ultimately, this will enable platforms to retrieve not only information but also meaning. RDF is based around the simple structure:

```
<subject> <predicate> <object>
```

where the subject and the object are the two resources to be considered and the predicate, the attribute that links them. The combination of the three is called a triple. Subjects are either a URI pointing to a resource or a blank node, while predicates are defined by URI and objects are either a URI or a literal like a number, a string, a Boolean, a date etc. When combined together, RDF triples draw a network of interconnected entities also called an RDF graph (example in Figure 2-13). An RDF graph and its set of triples are analogue to an ontology and its statements and can, therefore, be the foundation for ontological development in computer sciences. The original syntax for RDF graph was the XML syntax, also called RDF/XML. Nevertheless, other formats have been developed such as Turtle, JSON-LD or RDFa (for HTML and XML embedding) [300].

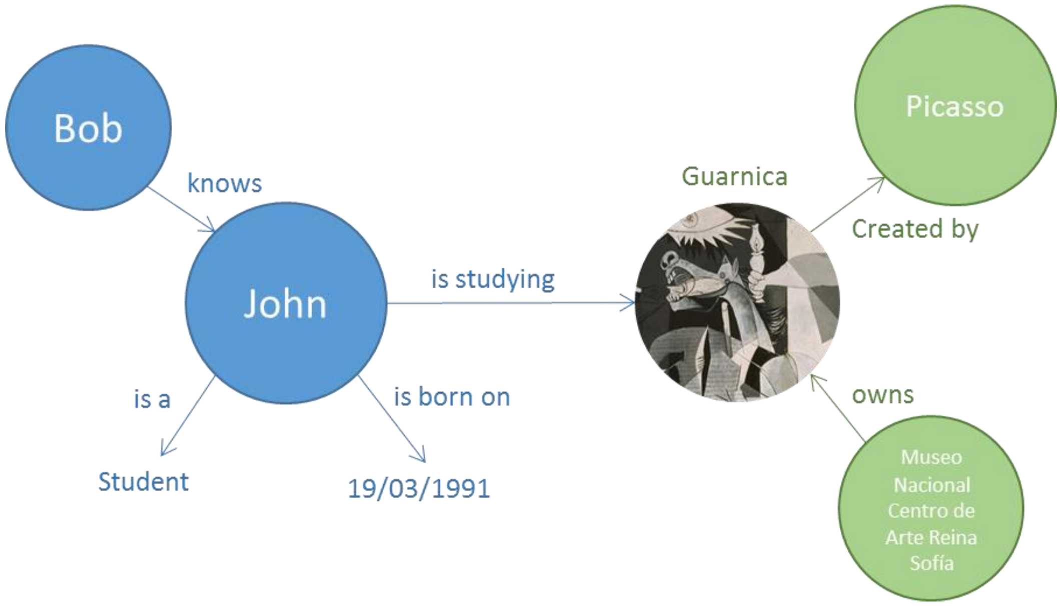


Figure 2-13 Informal graph representation

RDF is extended by the RDF Schema (or RDFS) which is a data-modelling vocabulary for RDF data [301]. This semantic extension defines concepts that group related RDF data and describe relationships. Table 2-4 gives the constructs that extended RDF with RDFS.

Table 2-4 RDFS extensions [32]

Construct	Syntactic form	Description
Class (a class)	rdfs:Class	A resource representing the class of all classes
subClassOf (a property)	rdfs:subClassOf	Predicate to define a subject as a subclass of an object
subPropertyOf(a property)	rdfs:subPropertyOf	Predicate to define a subject as a subproperty of an object
domain (a property)	rdfs:domain	Predicate to link a property with a class that is its domain
range (a property)	rdfs:range	Predicate to link a property with a class that is its range

The integration of those allows the development of more complex systems where information about the resources can be inferred. Additionally, RDFS gives the possibility to add comments, labels, references and definition to the RDF graph.

If the RDFS layer extends the RDF layer giving it a higher level of semantic, it still does not meet the requirement for the creation of an ontology that represents real-world assumptions [301]. It is missing essential features such as the possibility to assert characteristics to properties (e.g. inverse, transitive, symmetric), restrictions (e.g. range restrictions, cardinality, existence or strict distinction between classes and assertions) or relation between classes (e.g. disjointness, union, intersection), etc.

2.3.2.2 *OWL ontologies*

The Web Ontology Language (OWL) has been introduced in 2004, extending the RDFS functionality in order to fully support the creation of ontologies [302]. It comes in three different dialects OWL Full, OWL DL and OWL Lite. Figure 2-14 shows how the different dialects relate one to another. The OWL Full is the union between the OWL syntax and RDF. If it gains in expressivity, the OWL Full is considered to be to be undecidable to reason on [303]. Reasoning is an essential characteristic to take into account in ontological modelling. Indeed, an ontology is a logical system that is characterised by its capacity to reason and to infer implicit knowledge from the formally described knowledge base [304]. Conjunction, disjunction, equivalence, universal or existential quantifications and the possibility to construct subconcept/superconcept relationships (subsumptions) are the building blocks of such intelligent system in OWL. For a given graph, the dialect chosen to interpret the axioms will infer information in a different way following the extent it “bound” the language expressiveness. The reasoning process is influenced by the multiplicity of interpretations. If reasoning is open to too many interpretations, the dialect is untractable; if it is open to infinite interpretations, it is undecidable.

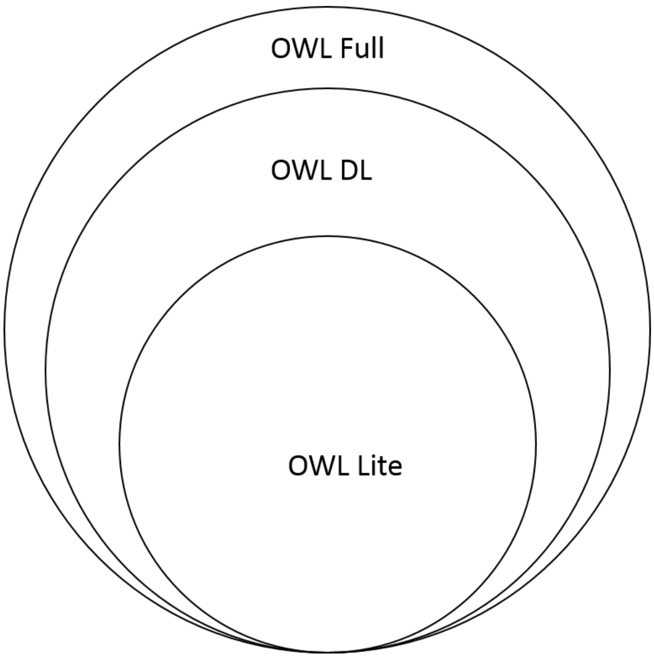


Figure 2-14 OWL dialects

Consequently, when it comes to reason over an OWL Full model, because the OWL Full does not include certain restrictions, it is harder if not impossible to compute. In other words, OWL Full is too expressive to meet computational requirements. Consequently, the OWL DL fragment has been developed as well. OWL DL ensures computational tractability by following the Description Logics semantic and adding restrictions that "bound" the expressiveness to a reasonable extent. It prohibits the use of classes as instances and distinguishes object and data properties. Finally, OWL Lite is itself a fragment of OWL DL that will support minimum requirements in term of expressiveness that includes cardinality 0/1 only, inverse, transitive and symmetric properties or existential (some values from) and universal (all values from) restrictions.

In 2009, a revised version of OWL, OWL 2 has been developed by the W3C [305]. OWL 2 is overall quite similar to OWL 1 but integrates features that are much needed for applicative purposes namely keys, property chains, richer datatypes and ranges, qualified cardinality restrictions, asymmetric, reflexive and disjoint properties, and enhanced annotation capabilities [305].

Similarly to OWL 1, OWL 2 integrates a DL fragment that follows the same principles as previously cited. However, as presented in Figure 2-15 OWL Lite has been changed by the creation of three distinct fragments or profiles that are OWL 2 EL, OWL 2 QL and OWL 2 RL [306]. Those three profiles are purely based on applicative perspective and are intended to be chosen following the reasoning tasks at hand. Indeed, following use to make of an ontology, it is interesting to be able to scale down the language expressiveness in order to meet reasonable computational requirements. OWL 2 EL is useful when it comes to reason over ontologies with a large conceptual part allowing reasoning in polynomial time. The fragment trims down some constructs such as universal quantification, cardinality restriction, disjunction and union as class constructors (unionOf, disjointUnion, dataUnion) or irreflexive, inverse, symmetric or asymmetric properties. OWL 2 QL is a fragment that deals with a large number of instances. It allows the query and reasoning over large datasets in Logspace with respect to the size of the data. Some of the restrictions in OWL 2 QL include existential quantification, enumeration of individuals and literals, keys, individual equality assertions or negative property assertions. Finally, the OWL 2 RL fragment is intended for uses that require a scalable reasoning without trading too much expressivity like in OWL 2 EL and OWL 2 QL. Here, all axioms defined by OWL 2 DL are present apart from unions as class constructor and reflexive object property axioms. The

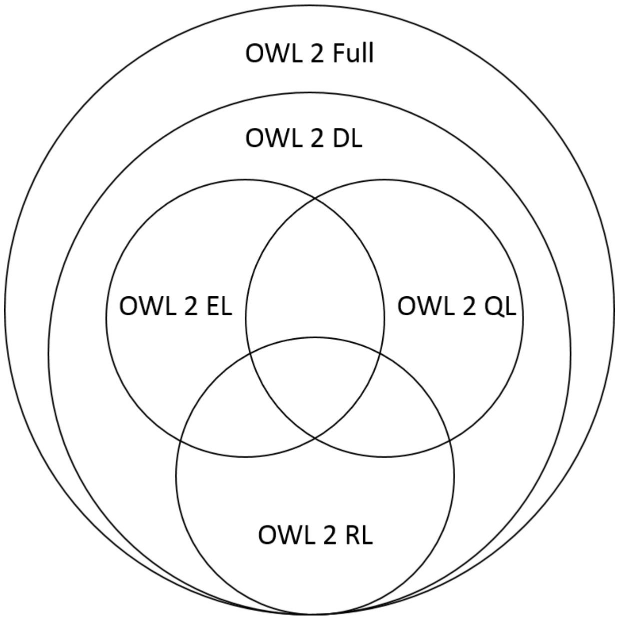


Figure 2-15 OWL 2 Profiles

difference with OWL 2 DL lays in some additional restrictions on the use of construct to specific syntactic positions [306].

2.3.2.3 *SPARQL*

The SPARQL protocol and language provides the possibility to query and manipulate RDF graph data in a similar way that SQL allows to query and manipulate relational databases data [307]. Four different types of query methods are possible in SPARQL: SELECT that retrieve resources matching the query; CONSTRUCT that returns an RDF graph formulated within the query; ASK that returns a boolean if the query matches a certain statement; and DESCRIBE that provides an RDF graph describing the resources interrogated. When querying, the clause WHERE enables to provide a graph pattern that one wants to interrogate against the existing RDF graph. Similarly to SQL, SPARQL integrates a set of constructs that allow aggregation (GROUP BY, HAVING), sequence and modifier (ORDER BY, OFFSET or LIMIT), algebra (SUM, AVG, MIN, MAX etc), string manipulation (CONCAT, REGEX, SUBSTR etc) and others. The SPARQL specification also describes how a SPARQL query must be evaluated against the different entailment regimes previously listed [308]. Indeed, a SPARQL query must be developed with regards to an entailment regime as the answer will differ from one regime to another [309].

2.3.2.4 *Future Perspectives*

To summarise, in the past 15 years, many efforts have been done for the standardisation and application of semantic web technologies especially on RDF, RDFS, OWL and SPARQL; the end goal being the creation of linked data and a unified global system of information. Despite interesting developments and its potential, the Semantic Web remains a niche that has a hard time to take off as a mainstream standard [310]. There are still a number of challenges that need to be overcome for the Semantic Web to take off [311]: (1) the ontologies themselves must be developed with more rigour and consistency and must be flexible enough to be updatable; (2) the predominant language used is English and efforts must be done for a multilinguistic system; (3) trust and proof mechanisms must be integrated to ensure data credibility and privacy; (4) it must be scalable to meet future requirements of a large scale implementation, especially in term of storage and reasoning power; (5) security must be re-enforced to ensure complete data privacy and protection; (6) and the usability must be improved for both users and developers. However, experts estimate that Semantic web will continue to grow as it is a major enabler for artificial intelligence, machine learning, and data interoperability [310].

2.3.3 Semantic for Smart Cities

Section 2.2 demonstrates how the Smart city vision lays on various cutting edge technologies such as IoT, BIM or GIS. In this section, the focus is on the integration of semantic web technologies within the smart city paradigm. Efforts for IoT, BIM or GIS semantisation will be reviewed and presented as a valuable solution to meet future requirements.

2.3.3.1 *IoT semantisation*

The interoperability issue is plainly specified within the IoT development agenda. With a focus on data interoperability, several solutions have been imagined as presented in Table 2-5.

Table 2-5 IoT interoperability solutions

Technologies	Description	sources
GSN	Global Sensor Networks is a middleware solution for sensor networks that introduced the concept of a virtual sensor. The virtual sensor is a key abstraction in XML format that provides all required information about the data such as meta-data for identification and discovery, data structure, SQL-based specification, functional properties on persistence, error handling, life-cycle management, and physical deployment. The virtual sensors then communicate in peer-to-peer using configurable wrappers.	[312]

XGSN	XGSN uses the existing capabilities of the GSN middleware and extends it by enriching virtual sensor with semantic using the Semantic Sensor Network (SSN) ontology described later in this section.	[313]
RFID Ecosystem	The RFID Ecosystem makes full use of RFID tags to transform low-level information into meaningful ones. They generate one tag-read event (TRE) per tag, antenna and time. The TREs go along a Tag Manager tool that allows one to define some metadata associated; a Place Manager tool that allows to precisely define location; and the Scenic Manager tool that connect RFID with related applications.	[314]
UBIWARE	UBIWARE is a middleware designed to tackle data heterogeneity by abstracting every entity as a data service using a REST/HTTP-based protocol. Data can be queried and manipulated via an API where each data service is identified by a URI and that defines PUT, GET, DELETE, UPDATE, INSERT and CALL operators. Additionally, a Data Service Manager (DSManager) exists that clusters different data service for subsequent management modules.	[315]
SOA-based	Devices interoperate by a service-oriented approach where parallels services offer devices functionalities as well as discovery and the inclusion of new functionalities. The platform layering architecture allows different level of services with the inclusion of Service mapper, Service Implementation Repository and Device Service Proxy that ensure that composed services run efficiently.	[314]
ROS	Robot Operating System has been used to unify heterogeneous sensors, actuators and services. Its architecture involves decentralised peer-to-peer, network concepts, publish-subscribe information distribution or bi-directional services between components. It is organised around nodes which are the processes, messages between nodes, topics which are attended application and services that pair topic with request and response messages.	[316], [317]
OGC's SensorThings API	The OGC's Sensor Things API is based on Sensor Web Enablement (SWE) standards that have been developed by the OGC with the objective to describe static and dynamic data and their meta-information. It includes the Sensor Model Language or SensorML, a language that enables meta-information annotations about processes and processing associated with measurements and post-treatment of measurements. The O&M data model is equally important and will be covered later in this section.	[318]–[321]

When it comes to semantic web technologies for IoT semantisation, some frameworks have become predominant.

Firstly, the SSN ontology has been developed by the W3C Semantic Sensor Network Incubator group in 2011 in order to describe sensor networks capabilities and properties as well as the more abstract act of sensing and the result of observations [322], [323]. The ontology is based on 41 concepts and 39 properties of the Dolce Ultra-Lite (DUL) upper-level ontology that catches natural languages and human common-sense concepts [324], [325]. Some key concepts such as Units of measurement, locations, hierarchies of sensor types, and feature and property hierarchies were voluntarily left aside as some referenced ontologies already exist for those aspects. However, the SSN ontology specification gives alignment examples with those ontologies. In 2017, a new version of the SSN ontology has been approved, splitting the framework into two distinct ontologies: SSN and SOSA (Sensor, Observation, Sample, and Actuator) [326]. Figure 2-16 shows an overview of the SSN ontology concepts with a focus on the observation perspective. In this ontology, observations are the product of a sensing method that is realised by a sensor. A sensor here is an abstract notion that applies to everything that realises observation e.g. human being, computers, physical sensors. Physical sensors are then a possible instance of sosa:Sensor which is a subclass of ssn:System. Here, sosa:Observation is defined by its relationship with a sosa:ObservableProperty (e.g. temperature) and a sosa:FeatureOfInterest (e.g. air of a specific room). Moreover, each sosa:Observation are associated with their results that can take the form of a direct rdfs:Literal value or a sosa:Result entity for further semantics descriptions. Similarly, the time is associated with each observation in two possible ways: a simple XML time value (xsd:dateTime) or an entity derived from the Time ontology [327]. Additionally, the ontology allows the definition of the sensing methods used as well as metadata about the sensors network and deployment procedure. Those concepts applied to sensors and observations are equally applicable to actuators and actuations within the SSN/SOSA ontology.

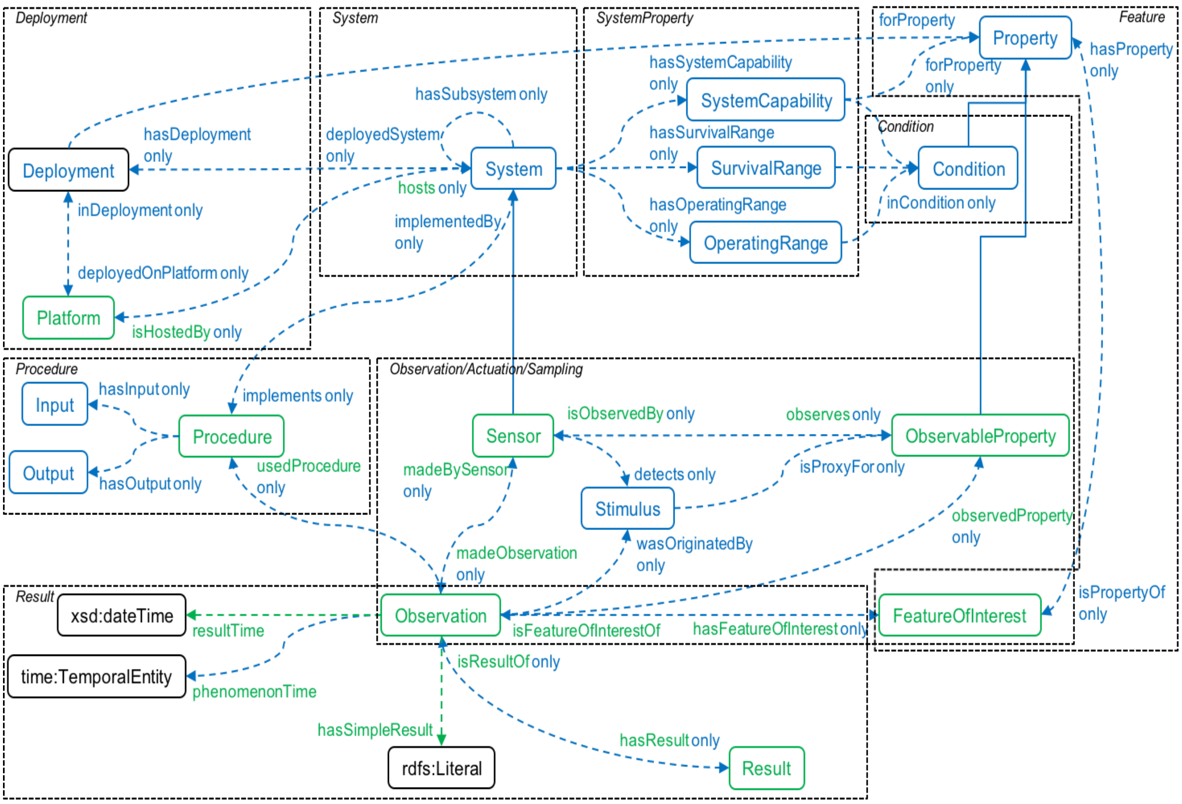


Figure 2-16 Overview of the SSN classes and properties (observation perspective) [326].

Another ontology to describe sensor networks is the SAREF ontology, standing for "Smart Appliances Reference Ontology"[328]. SAREF has emerged from an effort of the linked building data community for the standardisation of smart appliances semantics and is now an official ETSI standard [329]. The SAREF ontology is the outcome of the study of assets in the smart appliances domain such as projects, sets of documents, standards, working groups, committees, papers, or web pages [330]. Overall, 23 different assets have been used for the creation of SAREF in a bottom-up fashion through concepts mappings and alignments. The ontology focuses on three main concrete domains: (1) devices, sensors and their specification in terms of functions, states and services; (2) energy consumption/production information and profiles to optimize energy efficiency; and (3) building related semantic models. Figure 2-17 gives a sample of the SAREF design, centred on the class saref:Device. In SAREF, a device is defined as:

> "a *tangible object designed to accomplish a particular task in households, common public buildings or offices. In order to accomplish this task, the device performs one or more functions". Examples of devices are a light switch, a temperature sensor, an energy meter, a washing machine."* [330]

A device is then associated with different possible functions within actuation functions, sensing functions, metering functions or event functions. Those functions subsequently define what is being measured by the device and the time it has been recorded. Additionally, the device is associated with some meta-information about its specification as well as its absolute and relative location in a building.

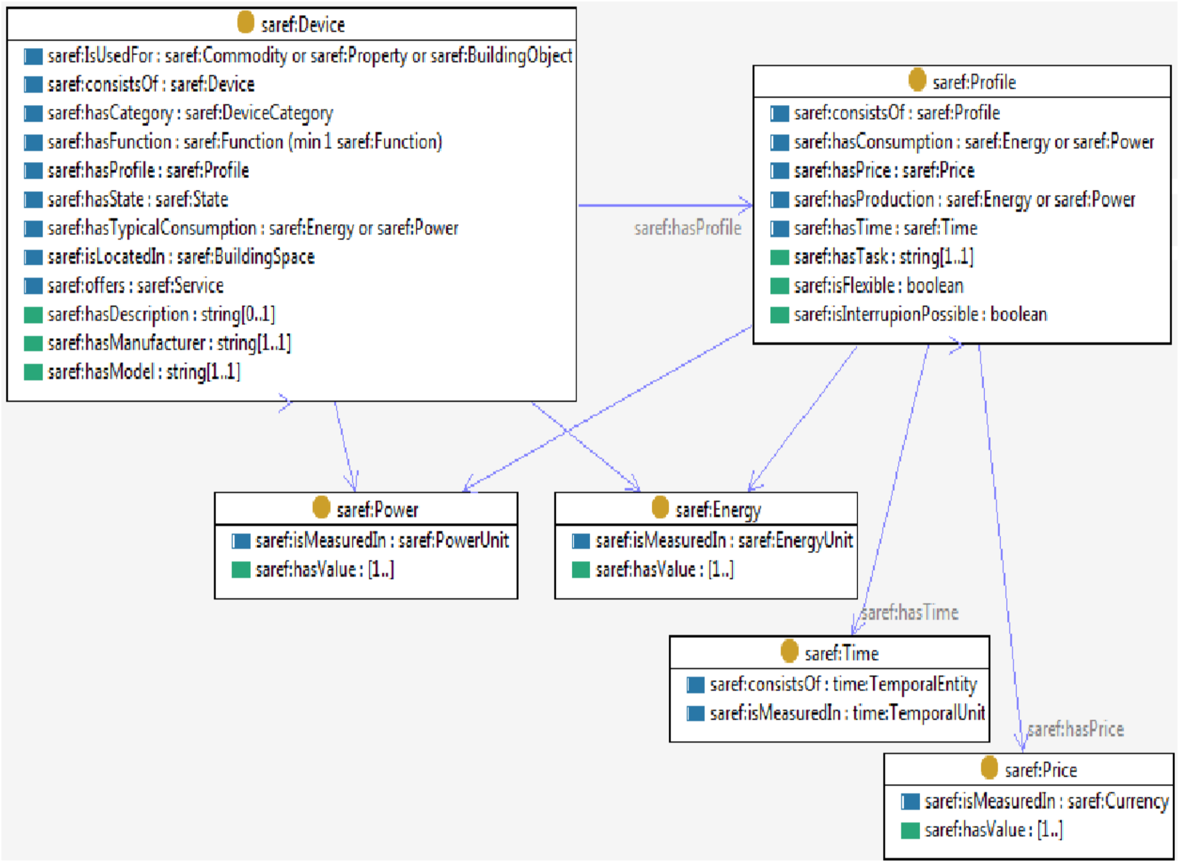


Figure 2-17 Sample of the SAREF ontology design, centred around the device concept [330]

Over the years, SSN and SAREF have been used in multiple projects and application that are listed in Table 2-6.

Table 2-6 SSN and SAREF use cases

Ontology	Description	Reference
SSN	Siemens have considered the ontology as a guideline for the creation of their Siemens ontology	[331]
SSN	OpenIoT, a middleware platform for the Internet of Things discovery created under the FP7 EU OpenIoT Project, uses it to represent their concepts	[28]
SSN	The SSN ontology has been used for the transformation of Meteorological Data into Linked Data [332]	[332]
SAREF	The SAREF ontology has been used within the H2020 EU THERMOSS Project in order to deal with interoperability within a collaborative network	[282]

2.3.3.2 *BIM semantisation*

The use of a standardised format to encode BIM models such as the IFC (Industry Foundation Classes) are believed to improve the issue of data interoperability between heterogeneous information systems [259].

The ISO 16739:2013 standard defines IFC as “*a conceptual data schema and an exchange file format for Building Information Model (BIM) data*” [333]. This open data model allows the transfer between platforms of different encoded construction entities like building components’ geometries and their physical properties [160]. The format is currently being promoted by buildingSMART that have pursued the works undertaken by the International Alliance for Interoperability (IAI) in 2004 [334]. Textual data, documents, images, numerical models or stakeholders annotation are a few examples of the type of data that can be related to IFC [160]. BIM software such as Revit proposes now an option to export their native BIM format to the IFC format. The option prefigures a growing interest in using the IFC as a common data format for BIM models. Lately, the growing interest in semantic web technologies has led to the creation of an alternative schema, the ifcOWL ontology [335].

IfcOWL is the preferred format to migrate from IFC format to semantic web technologies. In his literature review [335], Pauwels et al. advocate the use of ifcOWL to improve three main matters:

- *Interoperability*: ifcOWL allows a better information exchange and a common model understandable by machines and humans.
- *Linkage across domain*: ifcOWL unifies different domains such as GIS, energy, sensor data etc in a common model
- *Logical inference and proofs*: ifcOWL via its axioms can check consistency, regulation compliance or be used for the discovery of additional characteristics such as energy performance, costs, environmental impact etc.

Anumba et al. add that the shift to a semantic web environment could [336]:

- Reduce the amount of manual work (and thus mistakes) for terminologies translation, databases migration or standard conversion.
- Easily update data throughout the project teams and life.

The instantiation of the ifcOWL ontology can be done from scratch with the help of some ontology editor such as Protégé, Jena or the OWL API for Java; or it can be automatically instantiated via converting procedures. However, IFC files are often voluminous and complex and a manual instantiation is not recommended since it is time-consuming and can result in errors and inconsistencies [337]. Consequently, an automatic instantiation from an IFC-SPF file is favoured and many studies have focused on the creation of converting tools [337]–[342]. An open web-based convertor resulting from the work undertaken by Pauwels et al.[337] is also freely available [343]. However, the complexity and size of certain ifcOWL models instances greatly restrict an effective use. This issue is currently under investigation with the creation of a modular version where the ontology would be separate into subsets. Thus, users could load only the ifcOWL module that is useful for their application [344].

Overall, there are currently still some barriers that limit the adoption of BIM technologies with the AEC industry. Team division, resistance to change, technical problems, low level of training or business and financial issues are cited as factors that limit the use of BIM within the industry [258], [345], [346]. Additionally, legal, contractual and organisational concerns are mentioned [346], [347]. Despite those limitations, BIM adoption increases among AEC contractors, reaching a significant degree of maturity in Europe and North America (more than half uses BIM since 5 years or more) [348]. The sharp increase of BIM users in North America from 28% to 71% between 2007 and 2012 is representative of the current worldwide trend toward those technologies by the industry in the last decade [348]. This rise is driven by the recognized benefit of BIM on reducing errors, improving collaboration, enhancing organisation etc and overall a positive ROI. Finally, in their article [346], Rezgui et al. developed a BIM research and development roadmap that shows that BIM did not yet reach its full technological maturity and that the shift toward semantic web technologies, outsourced application processing, value proposition-based collaboration etc can trigger a new urge to the BIM adoption.

2.3.3.3 *GIS semantisation*

GIS has proven via numerous studies[188], [189], [203], [349]–[352] that it can rigorously catch interesting information about socio-economic insights of urban areas. Likewise the other features of the smart city paradigm, interoperability of GIS is essential in the construction of a ubiquitous information system.

Some standards have been developed in order to create Linked Data embedded with geospatial information. Indeed, the OGC has developed GeoSPARQL, a standard language that aims to support the production and query of geospatial data on the Semantic Web [353]. The GeoSPARQL ontology follows already existing OGC standards for its taxonomy to represent features, geometries, and their relationships [354]. It also provides spatial functions such as *ogcf:intersects* that can be used to query topological relationships between geographical entities. Moreover, already cited in the previous sections, the GeoNames ontology is often used within the semantic web community when it comes to describe geospatial information such as postal code, names, populations and relationships between labelled entities etc [355]. In contrary to GeoSPARQL that take the physical world perspective, GeoNames takes the social world perspective with the inclusion of social constructs that have been created by Humans such a countries, capitals, cities names etc.

Additionally, the OGC has conducted a series of activities for interoperability in relation with the smart city paradigm with the OGC Testbed 11 or the cityGML Quality Interoperability Experiment [356]. Already introduced in Section 2.2.2.2, cityGML is a data model, extension of GML, that supports the definition of 3D GIS [357]. GML is defined as:

> "*an XML encoding in compliance with ISO 19118 for the transport and storage of geographic information modelled in accordance with the conceptual modelling framework used in the ISO 19100 series of International Standards and including both the spatial and non-spatial properties of geographic features.*"[202]

This objective of such schema is to provide an open framework that can be used for geospatial application which will support storage and transport of geographic application schemas and the information they describe.

CityGML is, therefore, an application schema of GML that aims to model complex georeferenced 3D data along with their semantics [358]. CityGML is a valuable mean for 3D GIS to interoperate with BIM. Indeed, the integration of 3D BIM models within 3D GIS has been increasingly studied in the past few years. For example, Deng et al. have developed mapping rules between IFC and CityGML based on matching components for seamless domains interoperability [205]; Lin et al. have proposed an algorithm for the generation of 3D indoor route planning (a 3D GIS feature of the ESRI ArcScene software) from 3D BIM model formatted in IFC [359]; Yu and Teo have used 3D BIM models to produce four types of CityGML LoD models [265].

GML and cityGML being well-recognised open and free standards for the description of GIS, they have been used for the development of an ontology for the Semantic Web, the cityGML ontology. The ontology has been developed around the cityGML domains definition, including features such as terrain, land use objects, transportation, vegetation, water objects, sites (buildings, bridge, tunnel etc) and city furniture [360].

Figure 2-18 shows two features of the cityGML ontology directly inherited from the city GML schema that are the LoD for building and the transportation features [360], [361]. CityObject is an abstract object at the core of the ontology that encompasses the different domains related to urban areas (site, transport etc).

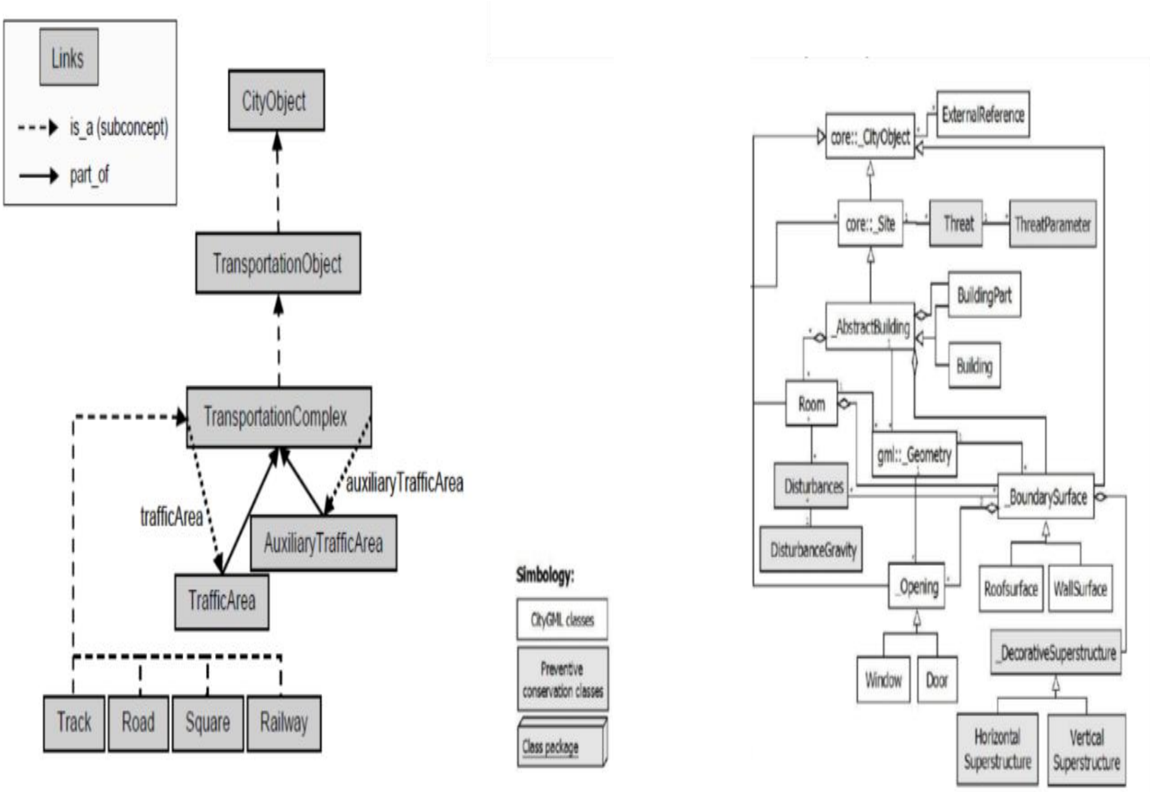


Figure 2-18 cityGML ontology transport and Building LoD sample [360], [361]

2.3.4 Urban Sustainability Related Ontologies and application

In the previous section, a certain number of ontologies have been cited as references for semantic web implementation of smart city components namely:

- SAREF ontology
- SSN ontology
- ifcOWL ontology
- GeoSPARQL ontology
- cityGML ontology

Those are believed to enable the semantic description of the IoT and their observation as well as buildings to urban scale features.

In addition to the ontologies previously cited, there are some essential domain ontology resources that can be reused to serve a more holistic view of the smart city and of the sustainable city vision. In the frame of the READY4SmartCities project FP7 project, efforts have been made for the creation of an ontology catalogue for smart cities and related domains [362], [363]. Table 2-7 takes advantages of this catalogue to present the most relevant ontological resources for the development of smart cities and urban sustainability-related ontology.

Table 2-7 Sustainable and Smart city related ontologies

Name	Urban Sustainability	Smart City	Smart Facilities	Energy	Water	Transport	Metrology	Description
Global City Indicator Foundation Ontology	✓			✓	✓	✓		The outcome of the PolisGnosis project [364] that semantically represented the ISO 37120, a standard with over 100 cities' quality of life and sustainability indicators [365].
OSMoSys ontology		✓	✓	✓	✓	✓		It describes several domains within the smart city paradigm such as energy, waste, water, transport, buildings etc and relates them to different data sources [366]
SESAME-S			✓	✓				The SEmantic SmArt Metering - Services for Energy Efficient Houses describes smart home components (meters, controllers, physical entities etc) as well as energy-related activities, events, pricing and policies.
SEMANCO ontology				✓				Enables the semantic representation of urban energy systems at different levels as well as energy-related data and related some energy simulation and assessment tools [367]
ee-district ontology				✓				Similarly, the ee-district ontology aims at linking energy-related data from local monitoring sources to high-level energy management software applications [368].
WISDOM					✓			WISDOM is a "Cyber-physical and social ontology of the water value chain" that unifies domestic water systems and their features [27]
Km4City						✓		The ontology defines in a global way all the part of a transport network form roads and transport-related furniture to administrative organisations [369]
Transport Disruption Ontology						✓		An ontology that catches travel and transport events to assess their disruptive impact [370]
QUDT							✓	Quantities, Units, Dimensions and Data Types or QUDT ontology supports any existing unit [371].

All those instances of ontologies that cover multiple features of the smart sustainable city can potentially be mapped on to another and integrated in a manner that it can create a holistic semantic representation.

2.3.5 Review of the Methodological Approaches for Ontological development

When it comes to develop an ontology, one must rigorously draw concepts and relationships that will truthfully fit the real world. It is a difficult task as, in order to build a global knowledge system, the developers must consider the already existing body of knowledge and how the newly developed ontology fits in it. Consequently, an ontology must position its degree of abstraction against its intended application. Indeed, it should be generic enough for potential reuses and extensions but specific enough as an over generalisation can leave relevant domain knowledge aside [372], [373]. For that reason, the semantic web community has developed a set of methodological guidelines that aims at rigorously setting the ontology domain scope (What? Where? When? How? Who?). This will ensure the development of well-design ontology that precisely defines its place within the existing models' landscape.

Several methodological frameworks exist in the literature for the development of ontologies. One of the first instance of methodological frameworks used is the Uschold-Gruninger methodology [373]–[375]. In a series of articles, they proposed a skeletal methodology for building ontologies, providing stages that can be used in any future methodology. Their methodology is centred on 4 mains steps:

1) Identify Purpose;
2) Building the Ontology;
    - Ontology capture,
    - Ontology coding,
    - Integrating Existing Ontologies;
3) Evaluation;
4) Documentation.

Figure 2-19 gives an overview of the different steps, starting with the purpose and scope identification; an essential stage that allows to better frame the ontology by identifying the intended uses and users. This can be done partially with the use of scenarios and competency questions. Indeed, one can imagine some scenarios related to the ontology purpose as the basis to form a set of questions that express reasoning problems that he/she wishes to solves [373]. Consequently, this set of questions formalises expressive and reasoning requirements that the ontology must fulfil. When drawing competency questions, it is advised to start with more general ones and to narrow them down to more specific ones [376]. From there, one can either start from scratch without formally describing requirements specifications, go for the development of an intermediate informal ontology or formalise the requirements to directly proceed to the formal ontology development. The authors advocate for a middle-out fashion design rather than top-down or bottom-up as they can lead to increase the level of details and efforts, the risk of inconsistencies and re-work needed. Overall, the ontology must prove clarity and be unambiguous, consistent and coherent, extensible and reusable.

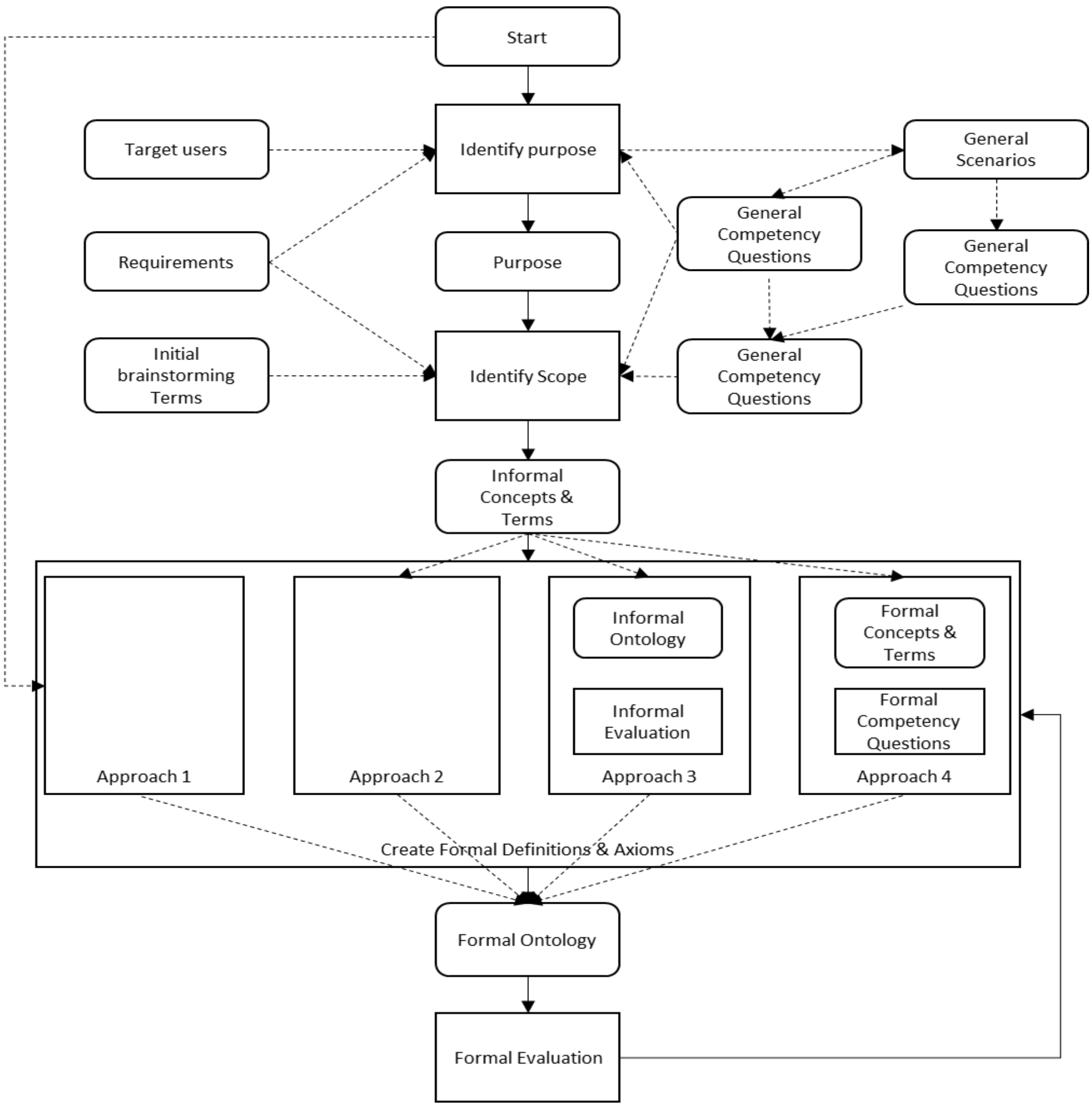


Figure 2-19 Uschold-King methodology taken from [373]

The METHONTOLOGY is a methodology to build ontologies from scratch, based on existing frameworks and experience of ontological in the domain of chemicals [377]. METHONTOLOGY borrows some core concepts from the Uschold-Gruninger methodology stages such as the requirement specifications, knowledge acquisition, ontology implementation and evaluation. However, METHONTOLOGY distinguishes itself from its parent by following a more structured and concrete approach. For instance, it gives guidelines for the elaboration of a first glossary, using any type of knowledge resources such as experts, books, figures, tables and other ontologies via text analysis, brainstorming, survey, formal or informal interviews etc. From this glossary, one must create different tables namely attributes, class attributes, conditions, rules, formulas, instances tables and form dictionaries of data and verbs. Additionally, they advocate for a complete reuse of existing ontologies along with an integration document that sums up the concepts reused and the ones expected but missing. Finally, they give procedures to evaluate, maintain and document the newly developed ontology so that it proves coherence and reusability.

One last instance of methodological frameworks is the NeOn Methodology framework [378]–[381]. NeOn is the most recent methodology and is inherited from the METHONTOLOGY. It has been designed as an extended version of the METHONTOLOGY that grasp in greater extent the specification of current ontologies, involving semantic web technologies such as RDF graph and OWL. This last methodology being used within the research, it will be described in more detail in Section 3.2.

2.4 SMART CITY VALUE CHAIN

Section 2.2 and 2.3 have introduced the Smart and Sustainable city concept from a technological perspective, listing a set of enabling technologies and how they can trigger changes. However, a more broad perspective is missing on the place of the different actors within this new paradigm and how to actually make it works. This section proposes an overview of the smart city value chain demonstrating the importance of citizens centred approach, collaboration and smart governance.

2.4.1 Smart and Sustainable city challenges

The smart city is believed to help the construction of sustainable places by the integration of increasingly accurate assessments and efficient services provision. However, if smart cities are strongly technologically based, broader issues must integrate their vision. Many challenges have been identified in the literature demonstrating that if the smart city is an enthusiastic perspective, it is also a real source of concerns [18], [125], [131], [155], [240]–[243]. In addition to the technical and technological challenges presented in section 2.2.3, socio-economic, managerial and ethical issues are involved.

2.4.1.1 *Economic & managerial issues*

The smart city paradigm is based on a large scale and involves many parties so that traditional business models are often not sufficient. Many new stakeholders come into play such as device suppliers, software suppliers, IoT operators, value-added services providers, data aggregators etc [382]. Consequently, IoT and Big Data from different sources can only achieve their full potential for the provision of efficient services via collaboration and communication between different entities [131]. A new vision of engagement models, rules, financing sources, and partnerships are much needed for the development and operation of smart cities [383]. Such concerns have already been investigated in the sector such as smart grid where collaborative networks and virtual enterprises are being considered for the creation of flexible, self-organized communities of intelligent and collaborative agents [146].

However, in the vision of a collaborative structure, some issues will emerge from potential tension and conflict between actors cause by ownership and commercial/competitive interests [384]. Strong leadership, ownership agreement and coordination are required to be well defined in order to overcome such tensions. Given the scale and nature of smart city projects, it seems that only governmental and public organisations are in the position to lead the adoption of the smart city [19]. In such prospect, public organisations must be careful of the partnership they built as, in a recent survey, 32% of business owners complained "*that cities tend to treat companies as suppliers or service providers rather than strategic partners*" [385]. It is important to better considered the role of every parties of a smart city project as a real partner that pro-actively serves common objectives.

It is certain that such a collaborative perspective will necessitate data exchange between partners. Thus, a challenging task already mentioned above relates to the combination of business, technical and data ownership [386]. Additionally, many concerns are upon sensitive data and data privacy which may discourage certain parties in taking part in such initiatives [386]. The development of data silos for value-added services by the public sector will, therefore, require the creation of arrangements for data management and sharing with private partners [22]. The value gained from open data access must outweigh the cost of acquisition and management of the data while preserving privacy and sensitive information [387].

To sum up, a new business model is required with the aim to better define the value of information and the services based upon, to encourage investments of private partners and to ensure privacy and fair competitiveness.

In spite of the significant investments, smart cities are not really taking off and not truly realizing the projected potentials [388]. The need for policy changes, scarcity of resources to trigger infrastructural investments, political uncertainties, lack of appropriate methodologies and metrics to monitor investment returns as well as the current sensitive economic context are as much barriers to the implementation and development of smart cities and Big Data [131], [388], [389]. Moreover, the ongoing smart city projects are seen as too "small" by the investors to consider scaling up the model to a larger scale [131], [388].

Another consideration should be toward human resources and skills requirement. Indeed, smart cities is still a young and emerging field which requires a large range of new skills [240]. Consequently, the democratisation of IoT and Big Data cannot be achieved without the training and provision of qualified workforces. This is especially true for the governmental workforce that, because of its core contribution, must have sufficient knowledge and expertise to treat with (maybe) more experimented private partners [384].

Moreover, the main features of Big Data from the IoT are variety, volume, velocity, veracity and value and its success lies on a large amount of valuable data collected from numerous heterogeneous sources. Great efforts must be done to ensure those features are respected especially data management methodologies to ensure data quality and semantics of voluminous datasets [390].

Overall, the domain is still at an early stage and not yet considered as self-sufficient to envisage fundamental changes. The increasing interest in the field by governmental institutions, scientific community and industries presages a situation that is improving.

2.4.1.2 *Ethical & social issues*

Big Data from the IoT involves the collection of data related to people. In order to have a better insight of the city, it is important to understand how people act and think. A key aspect of urban analytics is to profile the inhabitant of the city, understand their needs. To do so, many aspects are recorded such as incomes, health, activities, the degree of satisfaction toward services etc. Thus, profiling people cannot go without some ethical issues. If profiling citizen can be an extremely powerful mean to improve people quality of life, it can also be used at their expense for harmful purposes. There are then some important concerns on the privacy of the data [18], [131], [391]. The system can be hacked, giving valuable information to dishonest people [391]. They can be used by corporations for commercial purposes. Hollands in his article [392] argues that some underlying driving forces were on business-led, entrepreneurial or corporate urban development, creating what Michelle Provoost calls a "neo-liberal urban utopias" [393]. Data would then be the support of a market ideology, deviating from their primary goal which is to improve the human condition in the city. Finally, Big Data give to governments the possibility to significantly raise the level of surveillance in societies and individuals [18]. By doing so, societies are a step further to what Michel Foucault has defined as Panopticism [394]; model also presents in the successful novel *1984* by Georges Orwell [395] where a widespread surveillance would defeat the fundamental freedom rights of individuals.

Another issue is the development of a technocratic system of governance of the city [18]. In his book[396], Morozov warns against what he calls the techno-utopianism and solutionism; the idea that data and technology have the potential to solve any problems, even the more complex ones. Thus, the "city" would be perfectly understandable and we would be able to predict and respond with perfection accordingly [397]. This technocratic vision is reductionist and narrows the city to a superficial understanding excluding complex cultural, societal and political insights [18]. By giving too much importance to the data, we create a dependency on those ones and as a result to people that fully understand their operation, e.g. experts, corporation etc.

Overall, some argue that the inherent tension to systems that threaten privacy and freedom will inevitably grow with the growth of implementation of IoT and that the use of data will be a key problematic in the future societal, political and ethical landscape [18], [392].

2.4.2 Smart City Inclusive Development For Change

In order to answer the issues mentioned in the previous section, a certain number of measures must be taken that ensure that future smart cities are developed in people and environment best interest. This section will open on possible solutions investigated that aims for a more inclusive and fair vision.

2.4.2.1 *Collaborative Perspective and Open Data paradigm*

Figure 2-20 presents the layers stack of the smart city. Innovative approaches must be implemented at each of those layers in order to complete the smart city vision.

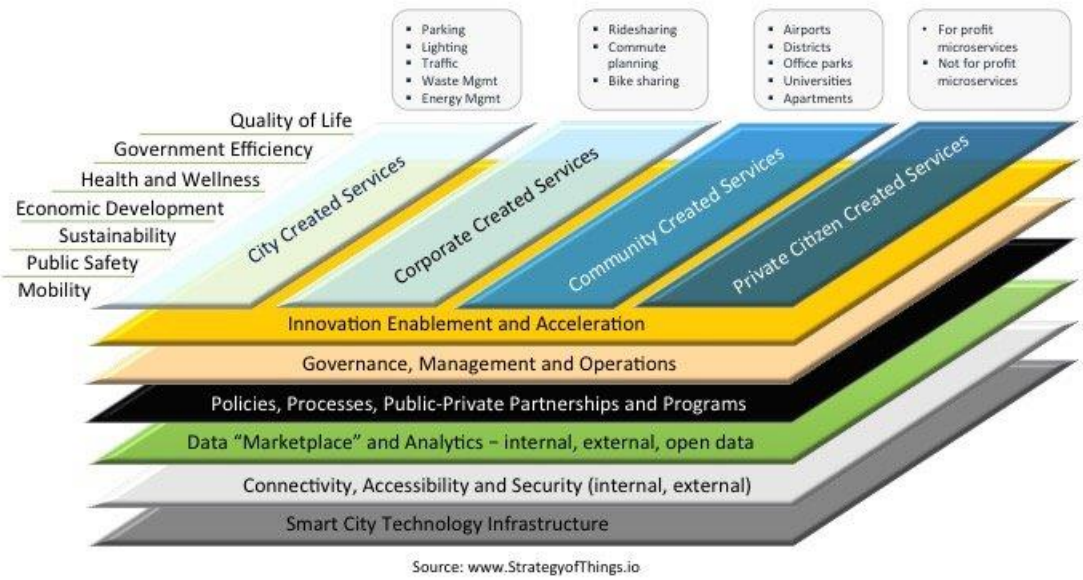


Figure 2-20 Smart city stack [383]

Starting by the **policies, processes, public-private partnerships and programs**, efforts must be done to frame the most rigorously possible collaboration between organisations [398]. This requires a new perspective on the actors' core contributions. Indeed, citizens should no longer be considered as passive consumers of services but become prosumer, actively engaging in shaping the city's future [385]. Initiative to include citizens in the assessment processes should flourish such as inviting them to invest in smart solutions, sponsoring or participating in hackathons and/or similar collaborative initiatives [385]. A survey of smart city business partners shows that 73% of executives agree in such participation of the citizen in the process. Same goes for engaging the private sector. They can no longer be considered as a simple service provider but should be fully included into long term public-private partnerships [399]. In that vision, collaborative networks (CNs) theory already mentioned in a previous section can be a valuable paradigm. Especially Virtual Organisation and Virtual Enterprises that are, by definition:

> "*a set of organisation/enterprises, public and/or private, that share resources and skills and have the potential and the will to cooperate via the establishment of a "base" long-term cooperation agreement and interoperable infrastructure. When a business opportunity is identified by one member (acting as a broker), a subset of these organizations can be selected and thus forming a VE/VO.*" [283].

Nevertheless, collaboration still rests on the "*potential and the will to cooperate*". Consequently, in order to convince the different parties in cooperating, efficient **governance, management and operation** must be achieved. Considering the nature of the smart city paradigm, governmental and public organisations should lead the smart city movement, engaging with the different actors and sponsoring initiatives. The political dimension especially is important. Governments must invest, provide institutionalized structures and policies to sustain smart city development [388], pushing developers to innovate and to provide relevant services to the citizens. Educational institutions accordingly with this politic must provide relevant training programs to produce skilled employees in this domain [240]. Good planning is important as well. All parties must agree on a development and management plan beforehand [398]. They must agree on practices and standardize the field to facilitate the exchange of knowledge and better provide services [398]. For instance, the systematic implementation of ICT in newly developed infrastructures. Good governance is the only solution to unlock smart cities full potential and to ensure the trust of all the parties involved.

There is a set of agreements and value propositions that governmental leaders must put forward in their governance strategy in the objective to ensure good cooperation between partners. Especially on key aspects such as the **data "Marketplace" and analytics – intern, extern and open data** and the **Connectivity, Accessibility and Security – internal, external.**

Those aspects are the prime source of concerns of citizens and private organisation as they can be harmful to their privacy and/or competitiveness. It relates directly to the openness of information and the willingness of people and organisation to share data. The strategy of open data is currently praised by research [18], [400]–[403] and governmental institution [404], [405] in the upcoming smart city paradigm. Opening data is believed to benefit people and businesses because they will open unexploited business and economic opportunities, improve sustainability and accelerate scientific progress and innovation (**Innovation Enablement and Acceleration**) [405]. More, the adoption of open data is motivated by its alignment with the ethical responsibilities of democratic governance for citizen participation, social inclusion in governance, and citizen empowerment [401]. However, this vision reflects two conflicting motivations, one being Open Data as a civic right, contributing to inclusivity and social progress, and the other being Open Data driven by predominantly economic and business justifications [401]. In regards to those tensions, governmental instances must introduce regulations and policies that will mitigate data monetization and capitalisation over open data. For instance, one solution suggests that data collected either by public only or PPP must be segmented into open and monetised data, where open represent the core and monetisation only concerns data for subsequent business related application [387]. In term of accessibility, regulation should be performed so that sensible and private information should be excluded from the open data vision and that ensure control and protection.

In such open perspective, it is important to give incentives to the private sector to invest in **smart city technology infrastructure** and services provision in cooperation with the public sector rather than deploying their own solutions and/or capitalizing on their data brokerage. A first approach should encourage initial investment with fiscal incentives and sharing the cost and risks of the investment. Then, in a second step, new performance-based approaches should define shared revenue mechanisms on the value of efficiency gains in service delivery, advertising generated incomes and revenue from value-added analytics services [406].

2.4.2.2 *Ethical Responsibility*

One essential aspect often mentioned in the previous section is the ethical responsibility the involved parties to ensure a fair and citizen-centred system. The ethical and social issues coupled with data should be studied in depth. Experts from different domain should come up with an ethical and legal framework that examine values and agendas underpinning data and whose interests they serve [18]. To this end, the European Commission, in 2012, proposed a new EU General Data Protection and Privacy Regulation (GDPR) based on the developed Privacy Impact Assessment (PIA) and Data Protection Impact Assessment (DPIA). Adopted in 2016 and implemented in 2018 [30], the regulation ensures good practices within the industry for personal information protection but is limited to the EU [245]. Additionally, some suggest that the ethical dimension of the domain should be part of the educational program for technicians [391].

Governance must prevent technological lock-in and technocratic development by relativizing the importance and power of data and by keeping the human dimension at the core of the system. As Jane Jacobs said:

"*Cities have the capability of providing something for everybody, only because, and only when, they are created by everybody*." [33].

2.4.3 Smart city redefined roles

2.4.3.1 *People*

In term of data collection, it is essential to apply a participatory approach where citizens can provide quality information in an efficient way [407]. This is essential because citizens have the ability to provide information that no other means can sense. Crowdsourcing information is a key aspect of the smart city paradigm as it unveils in greater details insight of the city metabolism[408]. Not only the citizens are the prime source of information but they are also the prime target of the services provided. The smart city is definitively shaped by and for the citizen. In such perspective, two types of participation can be distinguished, the "active" and "passive" participation [232]. In the passive participation, the data come from social media and all type of online inputs while the active participation solicits citizens to provide information about themselves via public platforms. In both cases, privacy must be respected and citizen must have sovereignty over the data they produced as it is the case in the Barcelona smart city movement [113]. They must have the power to withdraw their participation and personal information without conditions. It is the only democratic way to ensure citizen involvement in the smart city vision.

The smart city has the potential to amplify the Vox Populi and citizens must realise that they are no longer simple consumers but also actors of the cities operation [22]. In such perspective, data provision could appear as a civic initiative on the same level as voting.

Finally, the term “community” is often mentioned when speaking about citizens [22]. Indeed, governments will most likely deal with citizens aggregates when building up long term partnership. Segmenting the population into communities is important as it eases the interaction with subsets of common interests. Additionally, citizens integrated into communities have a greater feeling of belonging and thus a greater will to get involved in improving it. In such perspective, efforts have already been deployed to empower communities in the UK with the Localism Act 2011 [409]. This bill has been introduced to decentralise governance toward local governance. Local institutions are best suited to tackle certain matters due to their proximity with sites and residents. Allowing local authorities to act with more freedom improves the design and delivery of services. Additionally, a key aspect of such strategy is the increased participation of citizen in the decision process via the inclusion of new rights for citizen (right to challenge and right to bid). Those give the opportunity to citizen to collaborate with local governance on the core issues of that face their community.

2.4.3.2 *Businesses*

Businesses have a big role to play in the smart city development. Indeed it is most likely that cities won’t be able to finance all the projects and infrastructures required [410]. New collaboration forms between public and private organisations must be created with innovative financing models in order to build the smart city vision. Therefore, businesses should seek for opportunities to partner with cities and look beyond financial assets by considering shared returns on investment [411].

Moreover, the private sector is often the cradle for innovative approaches thanks to a more experimented workforce and a greater inclination for disruption. Helped by the public sector, businesses should drive innovation by for instance the inclusion of citizen participation via hackathons or other equivalent initiatives. Part of this vision includes the inclusion of open data as a factor for scientific progress and innovation.

Beyond the legal framework that goes along the smart city development concerning the gathering and uses of data, businesses have the ethical responsibility to protect personal information for any subsequent potentially harmful utilisation. This includes marketing and capitalising over information that at core does not belong to them.

2.4.3.3 *Government*

In the frame of the Smart city vision, it is interesting to look at government and the data they produce. Indeed, governments are the largest collector, holder and producer of citizens, organisations and public services related data [412]. This defines government as a major actor in the smart city landscape. Those information can relate to a more "private" aspect of people and organisation lives and therefore requires a trusting relationship between parties and a substantial effort on privacy protection [407]. The status of the government makes it a privileged partner for the exchange of such information because, as a non-profit and impartial institution, has no interest other than citizen protection and quality of life. However, this trusting relationship must be maintained and even strengthen. In that context, many advocate that government must open up on data collection and provision processes. Indeed, an open government is believed to improve citizens' trust towards the public administration by its strict policy on transparency [402]. In term of collection, various governmental participatory platforms already exists, to only cite a few : The NYC Participatory Budgeting allows users to share ideas for the annual participatory budget by pining them on a map [413]; PopVox gives the opportunity to citizen in voting for oncoming US Congress bills and legislation [413]; 311 Service Tracker Chicago lets people report malfunction or events that require public services assistance [414]. In term of data provision, some fundamental principles, present in Table 2-8, have been described by the Sunlight Foundation, a national, nonpartisan, non-profit organization that promote open data, policies and civic technologies [415].

Table 2-8 Sunlight Foundation ten principles for open government [415]

Principle	Description
Completeness	Datasets should be as complete as possible including raw data and metadata. At the exception of sensible data.
Primacy	Datasets should originate from primary sources and collection details must be provided.
Timeliness	Datasets should be realised as quickly as possible giving priority to time-sensitive information and real-time updates.
Ease of Physical and Electronic Access	Datasets should be easily accessible whether it is via physical or electronic ways
Machine Readability	Data stored should be in a machine-readable format for the machine to parse information.
Non-discrimination	Non-discriminatory access to data at any time without the need for identification or justification
Use of Commonly Owned Standards	Formatted data such as spreadsheet must be accessible via open sourced software and standards.
Licensing	The clear provision about the information being part of the public domain
Permanence	Data should be made available online in perpetuity.
Usage Costs	Data should be accessible at no cost or minimum

possible costs.

Those principles serve as a baseline for good practice in the implementation of open government, especially on how data should be present to the public in an accessible, transparent and reliable way. The rate at which local officials work with open data is growing and there is already numerous example of cities that share their data via online platforms, covering various aspect from the environment to art and culture or safety [416]–[418]. Note that Sunlight has initiated a benchmarking of 263 US cities based on their online data portal and the information they cover [419].

2.5 DECISION MAKING AND THE VALUE OF FORECASTING

Urban sustainability assessments aim to score the overall performance of a neighbourhood. To do so, it is important to foresee the impacts of certain decisions and actions taken within the neighbourhood. This is especially true when it comes to real-time assessment. In this section, decision support tools will be reviewed and compared to current urban sustainability assessment scheme. Arguments will be given for conceptual shift from simple assessment to decision support tool relating to the real-time nature of decisions taken in the operative stage of an urban area. Forecasting models applied to electrical load forecasts will be reviewed as an example of how forecasting is relevant in decision support and real-time assessment in general.

2.5.1 Overview of the Field

In the prospect of developing an assessment framework, it is interesting to look at what follows the assessment; what types of decisions will be taken. In fact, some even argue that the core aim of assessments is to facilitate decision [24], [420]. Decisions taken at the design or retrofitting stage of a development are likely to be more structural and in depth. They can potentially change the face of the neighbourhood. These decisions and the actions that follow are meant to last over a long time span. The impacts of those will highly influence the overall quality of the neighbourhood. Decisions taken at the operative stage are not design-related but are closer to day-to-day managerial decisions. The idea is to realise savings by an efficient and informed management of the services and facilities that the neighbourhood offers. Then the impact does not come directly from the individual actions themselves, which have a short time impact, but from their aggregation and iteration. This follows the observations of Figure 2-5 made in Section 2.1.2.3 that shows the impact of good operation and maintenance on building performance, extending its life span. A neighbourhood is not different and requires relevant operation and maintenance decisions in order to fulfil sustainable requirements.

The introduction of the near-real-time dimension of an assessment that can be applied at the operative stage of development constitute a fundamental shift in the vision of a sustainability assessment. Indeed, since decisions taken for new designs or retrofitting actions have long term impacts, they do not require sophisticated decision support systems; simulation, expert knowledge and assumptions are often sufficient. In most of the cases, the assessment does not go beyond its statue of simple scoring scheme since the decision and actions have already been taken. The new framework moves away from this vision since it is meant to be implemented in real-time and during the operative stage of the development where decisions have short term impacts. The global time frame of the scheme is dramatically narrower and assumptions and/or expert knowledge might be not enough in this situation. In this scope, the new framework must not only rate the performance of an urban area but also play the role of decision support system (DSS). It should consider both decisions with long term and short term impacts.

2.5.1.1 *The Role of Forecasting for decision support*

There are generally two different types of pattern that can be discovered when mining for information: explanatory that highlights interrelationships and affinities among the attributes; and predictive that foresees future values of attributes [421].

In 1980, Alter developed a taxonomy of DSS [422], introducing hierarchy between the different levels of decision support from simple file drawer systems to suggestion model, also referring data analysis systems, analysis information systems, accounting models, representational models and optimisation models. Although the taxonomy has been developed in the 1970s, it remains relevant and is still widely used within the DSS research community [423].

The taxonomy suggests that at a certain level of complexity (e.g. Representational models), decision support must estimate consequences of action, introducing the need for prediction and forecasting models.

Turban in 1995 [156] does also classify decision support models into seven categories: Complete enumeration – few alternatives, Optimisation via algorithm, Optimisation via analytical formula, Simulation, Heuristics, Other descriptive models and prescriptive models [156], [424]. Here again, can be found the notion of prediction.

Overall, decision support for business intelligence is a great example on how real-time data coupled with efficient data mining can improve the field of business and finance. It is particularly relevant to urban analytics since it leads to a better management of the urban environment in real-time.

2.5.1.2 *Current decision support systems for urban sustainability domains*

Many decision support systems (DSS) projects have been found in the literature toward urban metrics analysis. They all have been developed in relatively recent years which highlight the fact that it is still an emerging field.

ARIADNE is a decision support tool for efficient urban planning through life cycle analysis simulation developed by Les Mines Paris-Tech, France. The tool is partially based on EQUER, a life cycle analysis tool at the building level, and integrates 4 new indicators at the urban level. Forecasts are based on simulation and statistical regression assessing the impact of different potential urban planning on a long time span [425].

Similarly, the tool NEST (Neighbourhood Evaluation for Sustainable Territories) has been developed by the Centre de Ressources Technologiques Nobatek and the GRECAU laboratory of l'École Nationale d'Architecture et du Paysage de Bordeaux. This Sketch-Up plugging allows the user to visualise the environmental impacts of his urban design via a life cycle analysis simulation [426].

ODYSSEUS is an Open Dynamic System developed by a European consortium enabling the 'holistic energy management' in urban areas. It is designed to support operational and tactical energy decision making. In this scheme, external factors like weather forecast, energy tariff or peak demand are used for obtaining accurate prediction results through calculations and algorithms [427].

URB-GRAD is a Decision Support Tool for Retrofitting a District developed by a European consortium. A prediction enables the user to forecast and evaluate different scenarios for the district via advanced analytic and prediction algorithms [428].

ECODISTR-ICT project is an Integrated Decision Support System (IDSS) for sustainable retrofitting of urban districts with a focus on energy efficiency. The design team creates different alternatives which then will be assessed against KPIs [429].

Overall, if the tools have very different objectives, a better urban planning, energy efficiency, water efficiency, green design; they all share a common structure. Many of them use KPI in order to assess the performance of the area and does a comparison between a possible scenario, implemented by the user, and the current situation [429]–[432].

2.5.2 Forecasting model to support Urban Sustainability Assessment: A review.

Turban distinguishes two different types of forecasts: short-term forecasts (up to one year), that are used in deterministic models; and long-term forecasts that are used in both deterministic and probabilistic models [156], [424]. The most sophisticated forecasting methods are often introduced as a great support for “intelligence” in DSS, allowing the user to take more informed decisions, going beyond the simple expert knowledge.

In the cases of the DSS cited in Section 2.5.1.2, predictions are often the results of simulations or broad estimations of macro variables that influence at a big scale the overall neighbourhood (e.g. new urban planning). Other models go more into details, using advanced analytics and prediction algorithms. Odysseus project, for instance, focuses on energy saving in urban space during the operational stage of a neighbourhood life cycle [427]. It uses a prediction service to alert the end users in order to allow them to react and adapt their behaviour and enable forecasting of various energy-related scenarios. Urb-grade, a decision support tool for retrofitting a district, uses a similar approach by introducing a prediction module in order to extrapolate an output variable based on the trend marked [433].

It is interesting to look at the place of forecasting models within the current urban sustainability assessment frameworks mentioned in Section 2.1. The aim is to evaluate in which extent those schemes considered the use of prediction and what are the purposes of those. The authors have been looking in the frameworks’ handbooks the occurrence of the terms “forecast”, “prediction”, “estimation”, “simulation” and all possible declination of these terms. On the 29 tools, 11 contains some of these terms. There are 5 occurrences for “forecast-“ across the handbooks, followed by 17 occurrences for the term “simulat-“, 47 for the terms "predict-" and finally 55 for the terms "estimat-". Beyond the simple terms, the semantic has been studied. The high majority of the terms hold a broad meaning referring to general concept and/or organisational scheme (e.g. SWOT). Many terms are associated with macro variables such as climate change, climate hazards, and demography, contextualizing the current and future problems that cities face. In addition, when the prediction is associated with actual criteria such as energy, water, outdoor environment, economics, waste & materials, traffic, it often refers to what could be identified as baselines studies. It often does not refer to an accurate mathematical prediction but to a need for the stakeholder of a project to foresee the impact of their actions.

To sum up, forecasts in USA schemes are mainly on the long term and are often closer to assumptions rather than a clear methodical forecast. Frameworks being mainly developed for a design and development purpose, there are no specific needs for sophisticated and accurate forecasts. Assumptions and preventions are often enough for contextualisation and management.

2.5.3 Electrical Load Forecasting Models: a critical systematic review

Having highlighted the importance of predictive models within DSS and how they relate to the smart city paradigm and the newly developed USA framework, a systematic review of the available forecasting models have been done using the use case of electricity and power. The choice of electricity and power forecasts for this study has been motivated by their relevance in the smart grid field and by extension the smart city paradigm. In the prospect of efficient power delivery where multiple sources are in stake, prediction can ensure to foresees specific needs and therefore avoid peak loads and balance supply and demand [434]. Different types of prediction for power and electricity have been addressed in the literature from short-term, that are used for a rapid demand-response, to long term, to ensure capacity expansion, capital investment return studies, revenue analysis etc [435]. A multitude of models have been designed; to cite a few: multivariate regression [436]–[440], ANN [441]–[446], SVM [444], [447], [448], time series analysis [449]–[452]. However, there is still no consensus on the uses of the models for specific situations. The choice of a model is often left to the expert preference. Moreover, following the M3 competition (the 3rd occurrence of a series intended to evaluate and compare the accuracy of different forecasting models) Makridakis concludes:

"*simple methods developed by practicing forecasters do as well, or in many cases better, than sophisticated ones*"[453].

This prefigures that there are no evidences that sophisticated models will outperform simples ones and that when it comes to choose a model for a specific application, the doubt remains entire.

2.5.3.1 *Systematic review protocol*

In order to define which model is preferred by the research community for a particular case, a systematic literature review have been done. Over 120 different case study spread into 46 internationally distributed articles have been used for the comparison [454]. The research criteria for the section of resources were based on keywords such as "electricity forecasting models", "electricity prediction models" or "electricity demand models". With the use of Scopus, one of the established abstract and citation databases of reviewed literature [455], the occurrence of those terms have been searched within the threefold Title-Abstract-Keyword. Figure 2-21 gives the literature resources selection process. From the 10 667 original results, the articles have been narrowed down to 153 by selecting specific domains namely, "Engineering", "Energy", "Computer Science" and "Environmental Sciences", writing exclusively in English, comporting keywords such as "building", dwelling", "household", "energy demand", "electric load forecasting" and "energy demand". Among the 153 references, 46 have been reviewed in depth [434], [436], [445]–[448], [450]–[452], [456]–[458], [437], [459]–[468], [438], [469]–[478], [439], [479], [440]–[444]. The 41 references count around 113 different use cases of forecasting for both electricity and power.

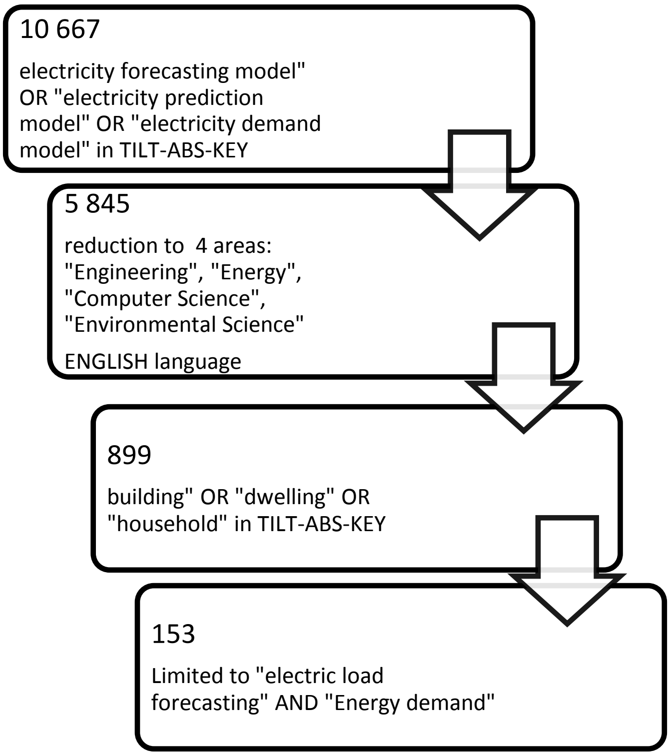


Figure 2-21 Selection procedure

2.5.3.2 *Forecasting models background*

All the 113 use cases have been compared against the characteristics listed in Table 2-9.

Table 2-9 Cases characteristics

Frame	Describe the context of the study and give an overview of the study purposes.
Location	Country of the case.
Scale	Scale of the study, from a single building to an entire country. Sizes of the sample are taken into account.
Term	From very short (1 min ahead) to very long term (several years ahead), gives the timeframe of the case.
Time resolution	Gives the time step considerate in the forecast: every minute, hours, day, years…
Inputs	Inputs implemented in the forecasting model.
Historical Data	Gives the length of the data sample used for the prediction as well as their origin (meters, statistical…).
Pre-processing	Indicates if the data have been pre-processed before being introduced into the forecasting model and which type of pre-process have been done.
Forecasting model	Gives the forecasting model employed

Error	Gives the type of error measured (CV(RMSE), MAPE, RMSE…).

Those characteristics have been selected after a first brief review of the references. They are the ones with the most occurrences. They appeared to be the most sufficient and relevant information to describe the particular settings of a use case.

Within the 113 applications, 16 different prediction models have been identified. Figure 2-22 gives the distribution of the different models in the papers reviewed.

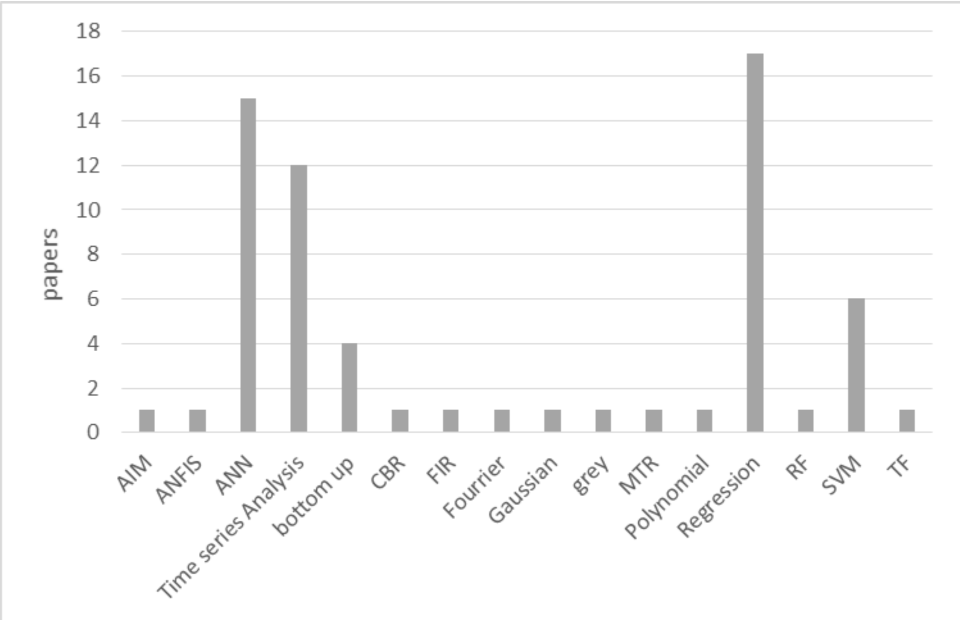


Figure 2-22 forecasting models distribution

Because one reference can proceed to several applications of a specific model, the distribution of the different models per paper is more representative of the actual trend. Note that “Time Series Analysis” covers the AR, MA, ARMA, ARIMA, seasonal or not, with or without exogenous variables. The observations show that some models are clearly favoured. The regression model covers multiple regressions or multivariate regressions. It is present in 17 papers out of 41, or 43.6% of them. This is followed by the Artificial Neural Networks (ANN) that can be found in 15 references or 38.5% of them. Then Time series are present in 12 references or 30.8% and followed by Support Vector Machine (SVM) and Bottom-Up models that are present in respectively 6 and 4 papers or 15.4% and 10.3% of them. Finally, the other models are all singularities. The relatively high amount of regression, ANN and time series is also due to the fact that they are often used in order to compare the performance of a newly developed model. This re-enforces their status of reference models in the research community

A good practice is the pre-processing of data input in the models. Many studies advocates for the use of pre-processing techniques in order to improve the accuracy of the model [436], [443], [452], particularly when considering machine learning algorithms [480]–[482]. 2/3 of the references mentioned a sort of data pre-processing in their modelling. Common techniques include filling missing values and smoothing noise, measurement of variables dependency and significance, clusterisation and classification and check for stationarity and seasonality. Mathematical and statistical tools can be used such as principal component analysis (PCA) or Pearson correlation (PCC) to check correlation, analysis of variance (ANOVA) or Kernel density estimation (KDE), for instance.

2.5.3.3 *Review key observations*

The characteristics listed in Table 2-9 have been studied and the results are displayed in Table 2-10.

Table 2-10 Parameter instances

Parameter	**Instances found in the literature**						
Scale	Building	District	City	Region	Country		
Term	Very Short (<1h)	Short (1h to several days)	Mid (1 month to a season)	Long (>year)			
Time resolution	Subhourly	Hourly	Daily	Weekly	Monthly	Seasonal	Annual
Inputs	None	Socio-economic	Environmental (e.g. weather)	Building	Time index		

Those parameters have been evaluated against the most encountered models, namely ANN, Time series, bottom up, regression and SVM. Figure 2-23 presents the use of forecasting models against the time horizon. Regression and bottom-up models within the references are most likely used for long term prediction. On the other hand, ANN and times series analysis with respectively 10 and 9 papers are mainly being used for short term prediction. Same for SVM that is used in 5 out of 7 papers for short term prediction.

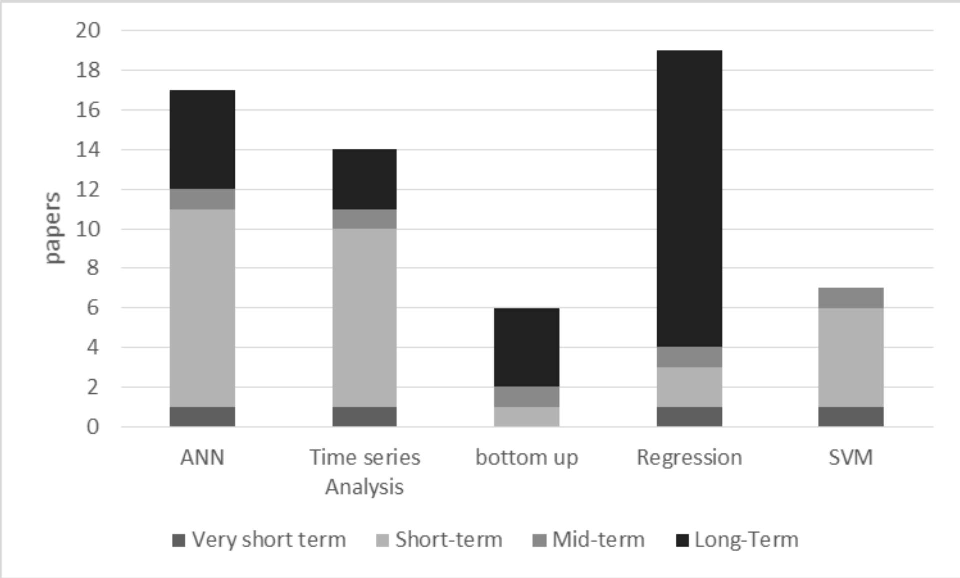


Figure 2-23 Models vs time horizon distribution

Figure 2-24 gives the number of occurrences in which variables are introduced in the different models. Regressions are mainly set up with socio-economic and building related variables. This can be explained by the fact the regressions are most often used for long term prediction and socio-economic and building related variables are the most significant in this configuration. ANNs seems to be mostly set up with environmental, building and time index inputs but also shows a relatively large range of possible inputs considered. This demonstrates a certain flexibility of the model against the variables introduced as inputs. Time series most often do not use exogenous variables. When introduced, those are used to reach a better performance which is not always guaranteed. Finally, bottom-up model systematically takes building related variables as it is by nature based on occupancy patterns and appliances related data.

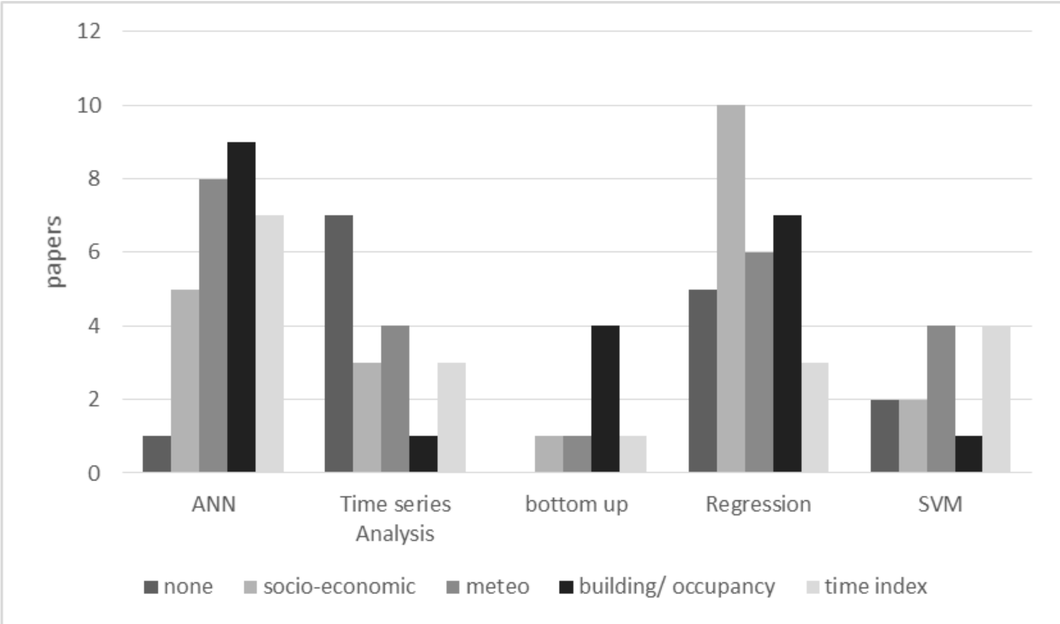


Figure 2-24 Models vs input distribution

Figure 2-25 show how each model is used in term of time horizon and resolution. ANN, Time series and SVM previously mentioned as being used for very short to short term prediction are most likely to use subhourly to hourly time resolution, rarely daily or more. Additionally, when it comes to long term prediction over several years, regression models seems to be favoured while bottom-up being by nature based on aggregation, a lower time resolution (hourly to weekly) is preferred.

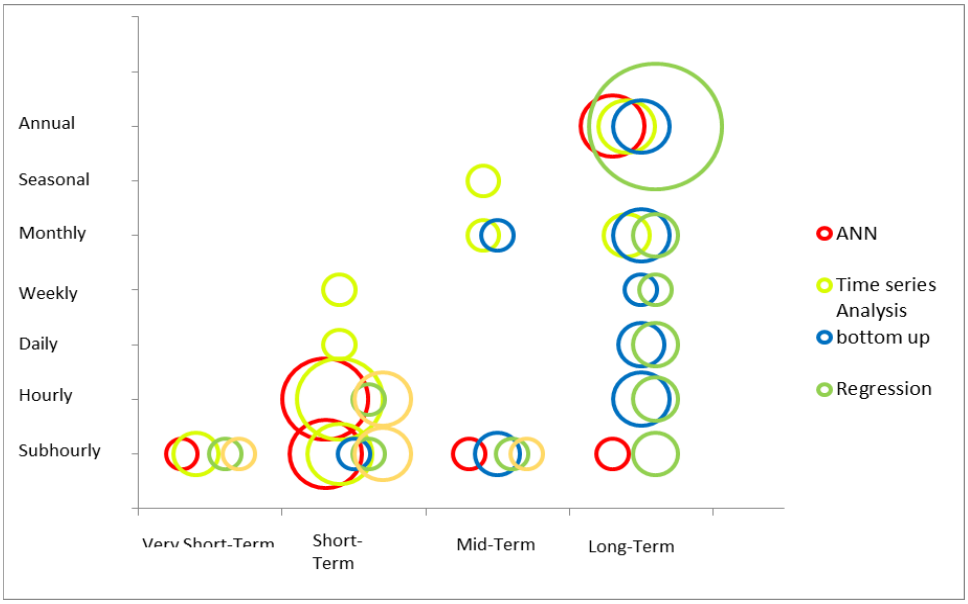


Figure 2-25 Models vs time horizon and resolution

The outcome of this systematic review has enabled the creation of a general taxonomy for electricity and power forecasting models selection. The user can refer to the taxonomy present in Figure 2-26 and chose the preferred model indicated by the colour code. Table 2-11 goes along the taxonomy, describing in detail the model.

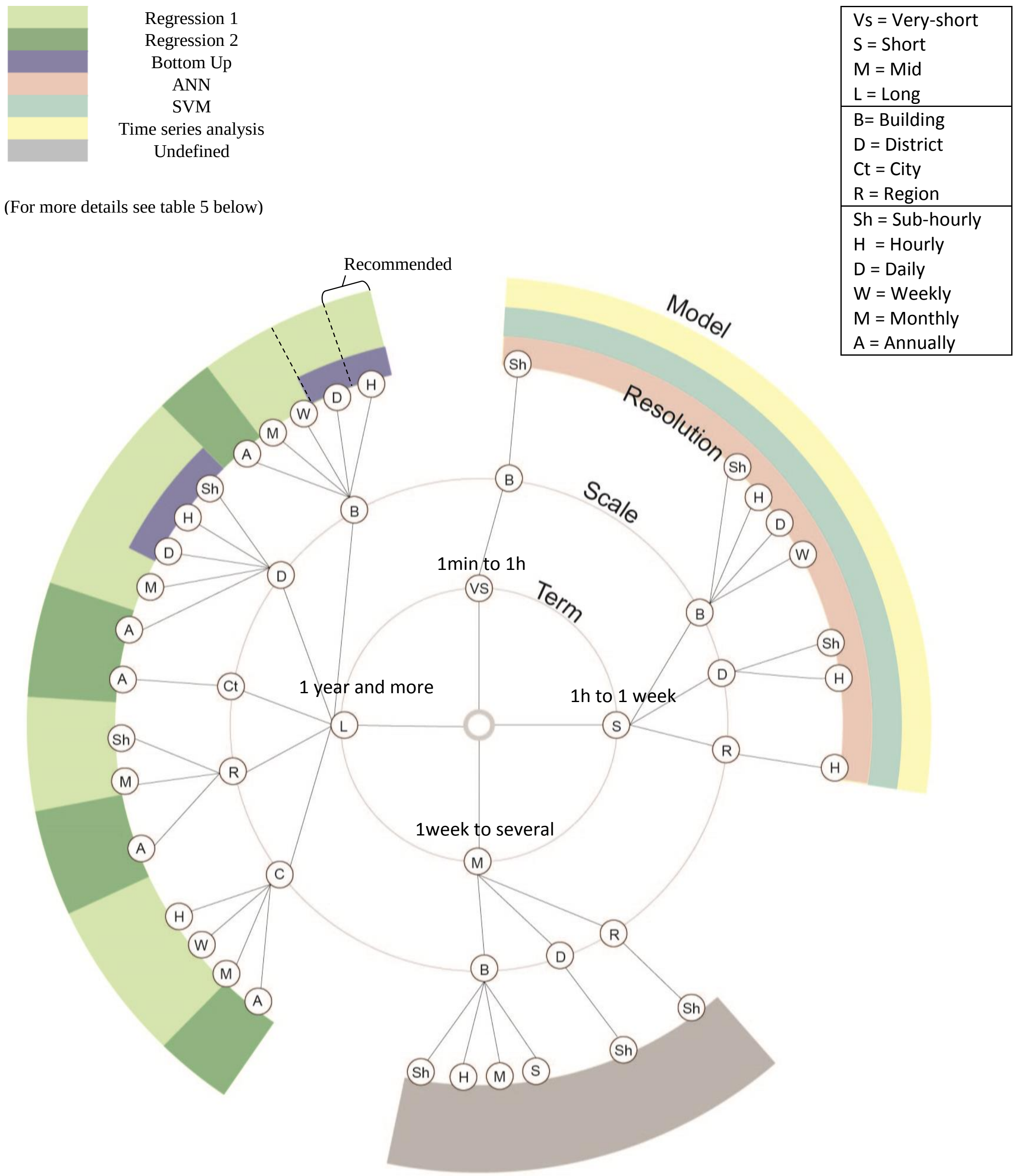


Figure 2-26 General Taxonomy

Table 2-11 General Taxonomy reference

Colour	Model	Inputs resolution	Inputs	Pre-process
	Regression	Subhourly to Hourly	Building data can eventually be introduced for better performance.	Smoothing high-resolution dataset is recommended. Clustering dataset in seasonal patterns can eventually be done to improve performance.
	Regression	Hourly to Annual	Socio-economic data are often introduced	Dependency & significance in order to lower the amount of input data
	Bottom Up	Subhourly	Building data are always introduced due to the nature of the model. Environmental data and time index can eventually be introduced for better performance.	Smoothing high-resolution data is recommended
	ANN	Subhourly to Hourly	Can be set with a large variety of data. Mainly environmental, building and time index. Time index often improves performance.	Dependency & significance in order to lower the amount of input data. Smoothing high-resolution data is recommended.
	SVM	Subhourly to Hourly	Can be set with a large variety of data. Mainly environmental, building and time index. Time index often improves performance.	Smoothing high-resolution data is recommended. Clustering dataset in seasonal patterns can eventually be done to improve performance.
	Time series analysis	Subhourly to Hourly	Environmental and time index data can eventually be introduced for better performance.	Smoothing high-resolution data is recommended. Clustering dataset in seasonal patterns can eventually be done to improve performance.
	The few amounts of cases on mid-term forecast does not allow any generalisation on the model to employ. A large variety of unique models have been employed in the literature.			

In conclusion, if the outcome of this study is valid for electricity and power, it must be extrapolated to other domains with care. The intention here was to give a good picture of the forecasting models landscape and how a certain domain expert community tends to use the most appropriated ones.

2.6 SUMMARY: THE USA FRAMEWORK, AN INNOVATIVE VIEW OF SUSTAINABILITY ASSESSMENT WITHIN THE SMART CITY PARADIGM

The smart city paradigm presents a unique possibility to improve sustainability and quality of life within urban areas. It is a really promising field, praised by public and private organisations as well as academics. Works have been undertaken in both hardware and software solutions for data collection, analysis, transmissions and the delivery of valuable services [125], [127], [136], [139], [161]. Some of the key aspects of the future smart city paradigm have been described in this literature review. The IoT is certainly the main enabler for the development of smart cities with the ability to consistently measure the city metabolism[132]. It will surely give great insights of the city operation. Additionally, technologies such as BIM or GIS have made significant progress in digitalizing real life information into ubiquitous, reliable information systems[176], [207], [254], [334]. Their inclusion into the smart city movement will certainly enable the creation of increasingly intelligent decision support tools where the digitalised city will meet real-world information and this in real-time.

In the light of such technological shift, the urban sustainability assessment schemes (see Table 2-2) found and reviewed in the literature seems to disregard the value of the IoT and the smart city. Indeed, they most often focus on newly designed neighbourhood, being static and expert based. The domain seems predominantly composed of consultancy schemes where expertise is given in an ad-hoc manner following a certain set of criteria. Moreover, the schemes studied have shown a lack of consensus, transparency and participation. There is no doubt that those frameworks are efficient in their specific purpose; however, it seems that there is still some room for improvement in the assessment of the environmental, social and economic quality of an urban area [14], [25], [47], [48]. Especially when it comes to assess neighbourhoods' efficiency over its entire operation. There are really few instances of sustainability assessment for neighbourhoods in operation. Yet, this feature is much needed as the qualification of a neighbourhood as "sustainable" during its design does not guarantee that it will keep the same level of quality over its entire lifetime. Providing a regular assessment throughout a neighbourhood entire life cycle is the only way to ensure it keeps up with an increasingly strict set of requirements. Consequently, the sustainability assessment of an urban area is believed to leverage on the smart city paradigm. Data produced by the IoT is a valuable source of information for the determination of urban sustainability KPIs. Such approach will allow a dynamic interpretation of the urban environment enabling a context-aware, near real-time decision making.

The development of such assessment is challenging. Some issues inherited from the smart city paradigm are particularly relevant when it comes to measure sustainability at the urban level. Indeed, sustainability is a broad concept that includes various domains and actors. It is more than likely that diverse information systems and organisation will be involved, forming a heterogeneous set of entities that the assessment will rest on [146], [388], [406], [410]. Methods should be integrated as to “unify” heterogeneous sources for the creation of a ubiquitous system. A solution investigated in the literature review is the semantic web technologies that present data as data on the internet, freeing them from any format and specification[303], [483]–[485]. Semantic modelling can solve interoperability, improving the delivery of reliable data and advanced software[143]. The next chapter presents the methodological approach used in the creation of an Urban Sustainability Assessment (USA) Framework. The research will implement a mixed strategy both positivist and interpretivist with the use of DELPHI surveys, Case studies, intensive reviews and action research.

Design and Methodology

3.1 INTRODUCTION

The effective realisation of a research project requires a rigorous methodology. The adoption of a methodology must be well defined before starting any development. Research methodology has been the subject of many studies in the past, leading to a well-defined framework where different research philosophies and methods are described [32], [486]–[489].

In this chapter, an overview of the different schools of thought for research methodology is given. The research approach chosen to answer the research questions is described in detail as well as the processes used to conduct the research.

3.2 RESEARCH PHILOSOPHY

The first step of conducting a research methodology is the understanding of certain research philosophies. Indeed, Guba & Lincoln [488] advocate that the question of paradigm supplants the questions of research methods. They note:

> "*Questions of method are secondary to questions of paradigm, which we define as the basic belief system or world view that guides the investigation, not only in choices of method but in ontologically and epistemologically fundamental ways.*"

The paradigms are the building blocks of a philosophical perspective. Or, as Filstead puts it, a paradigm is a [489]:

> "*set of interrelated assumptions about the social world which provides a philosophical and conceptual framework for the organized study of that world.*"

Therefore, the understanding of certain knowledge relies on some assumptions that define our perspective of the surrounding world. Those assumptions are categorized into 4 main groups presented in Table 3-1.

Table 3-1 Philosophical Assumptions after Creswell [32].

Assumptions	Question	Characteristics	Example
Ontological	What is the nature of reality?	Reflects the different ways of perceiving reality by individuals. In simple terms, objective or subjective perspectives.	The researcher must take into account the different perspectives developed in its findings.
Epistemological	How is knowledge gained? What is the relationship between the researcher and its research?	Classifies what constitute the knowledge or not. The knowledge possibilities, nature, sources and limitations are of utmost importance.	The researcher involves him or herself with the subject of its study.
Axiological	What is the role of personal values in the research?	Focuses on what the researcher values in its research and finding.	The researcher is aware and assesses the extent his or her values affect the findings.
Methodological	What are the processes, techniques and languages to conduct the research?	Uses inductive logic, contextualises the topic and designs the study	The research details the context of the investigation and iteratively revises his or her findings

Therefore, each philosophical assumptions encompasses a set of research philosophies

Figure 3-1 presents Saunder's Research onion that summarized the different philosophy, approaches, strategies, choices, time horizon and procedures that constitute a research methodology. Some philosophies are intrinsically linked to the epistemological assumptions while others linked to the ontological assumption.

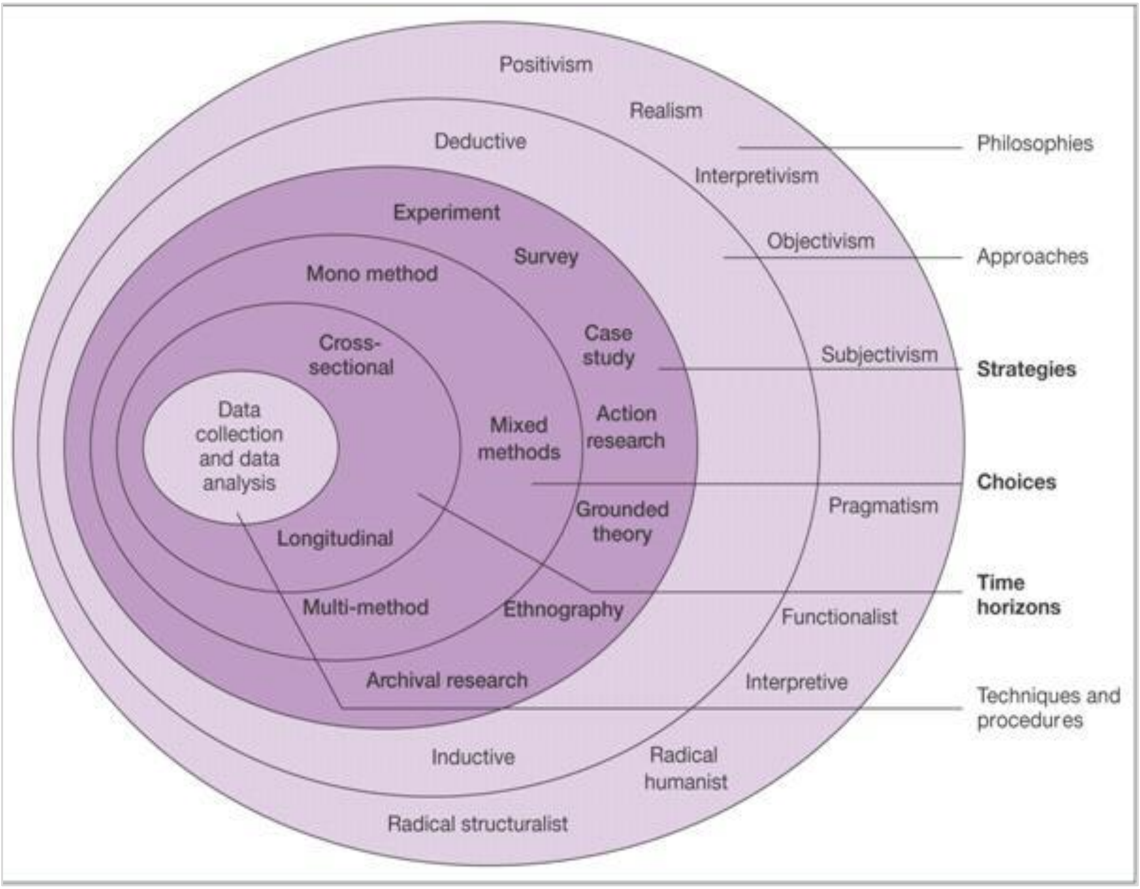


Figure 3-1 Saunder's Research Onion

3.2.1 Elements of Ontology

Ontology questions the nature of reality itself and the way it is perceived. There are two antagonist philosophies that aim at defining the notion of reality perception which are the objectivism and the subjectivism [490].

- *Objectivism* is an ontological position where social phenomenon and entities are independent of social actors [486], [487]. The social world is then a concrete structure where natural laws, patterns and relationships are considered as facts.
- *Subjectivism* (sometimes called *constructionism*) is an ontological position where social phenomenon and entities are only the result of the perceptions of the social actors. Therefore, they are constantly revised and the social world is not observed as definitive [486], [487].

3.2.2 Elements of Epistemology

Epistemology questions the conditions in which the knowledge is gained and what constitute knowledge. There are 4 distinct philosophical stances covered by epistemology:

- *Positivism* refers to the social world as an observable object where knowledge improves with observations and measurements. It is the stance of the natural scientist. The researcher conducts an objective approach, concerned about facts independently of his social perceptions. [491].
- *Realism,* similarly to positivism, undertakes a scientific approach to knowledge construction. It is the belief that real-world entities exist independently of the human mind. Subsequently, two types of realism exist, (1) the direct realism where what is being sensed is considered as truthful to reality; (2) the critical realism where it is acknowledged that what is being sensed may not picture the real world [491].
- *Interpretivism* relies on the idea that a certain reality exists only through the social constructions that are the culture, language, meanings etc. This philosophy relates more to social studies about human beings as social actors and has emerged after a critic of the positivism philosophy that could not catch the complexity of the social world [491].
- *Pragmatism* has the particularity to gather both positivist and interpretivist philosophies in the construction of knowledge. The pragmatic accepts that there is no single way to perceive reality and that multiple approaches are possible. Therefore, some advocate that the research question is of prime importance since it clarifies the mixed-position that the research adopts [491].

3.2.3 Research Philosophy Stance

The different research philosophies presented in the previous section must be seen as guidelines that enhance the way to conduct a research. Indeed, even if all different philosophies are well-defined, Saunders advocate that research philosophies are not intended to be chosen [487]. Instead, particular fields or activities are intrinsically embedded with certain philosophies. Therefore, a good understanding of the research philosophies is valuable in improving knowledge.

The research questions raised gives a good insight on both the technological and social aspects intimately link to the field of sustainability. Therefore, considering the previous definition of the research philosophies, it seems that the pragmatist approach would suit best to the research undertaken. Indeed, sustainability covers some aspects that are fundamentally related to “natural objects” subject to explicit laws and patterns, relying on quantitative data, hence, the need for a positivist stance. But sustainability equally covers aspects associated with social actors and how they perceive certain parts of the urban structures and social fabric. This relies on qualitative data and relates to an interpretivist approach. Moreover, the study questions the nature of sustainability itself, a domain that still did not reached consensus and that is subject to interpretations.

Moreover, the field of computer sciences is also overlapping the paradigm addressed in this research and its philosophical stance must be investigated. The interdisciplinary nature of computer sciences brings complexity to the philosophical stance to adopt as conflicting paradigms can emerge [492]. Additionally, computer programs and algorithms stand between an abstract representation of the world, coded within the lines, and a concrete application as it may execute concrete actions. Therefore, one must wonder where lay the value of a computer program, in its design or its application [493]. In the present case, the computer scientist approach is more related to the sciences of information and the interpretivist aspect that lay in the semantics and the value of ICT for sustainability assessment, as well as the software engineering that rely on empirical evidences such as the computation speed and its concrete function.

Consequently, in the light of the complex aspects of the domains, the "classical" choice between positivism and interpretivism is irrelevant and the philosophical perspective should be tailored to the research questions as the pragmatism suggests it. Table 3-2 gives an overview of how each research question relates to positivist and interpretivist philosophical stances, demonstrating the need of a pragmatic stance.

Table 3-2 Research questions against philosophical stance

Research questions	Positivism	Interpretivism
RQ1		Sustainability is not consensual and open to interpretation, especially on the inclusion of socio-economic aspects.
RQ2	• Certain physical phenomenons included in sustainability assessment are factual, observable objects in the real world. Decisions taken to improve them will trigger observable and quantifiable changes. • Optimisation and prediction accuracy can be measured.	Matters such as the log frequency of measurable aspects or the time horizon of predictive models are disputable. They relate to the time and the meaning of real time, a social construct.

RQ3	• ICTs integrated within the smart city paradigm observe and quantify phenomenon in the surrounding environment. • Faulty records can be detected and measurements accuracy can be quantified.	• Some models (e.g. BIM or GIS) are abstractions of the real world with a relative level of details. No matter their level of detail, they will always "only" be abstractions. • Other models supporting sustainability KPIs are by nature based on subjective interpretation (e.g. data mining of social media to measure satisfaction and feelings or thermal comfort)
RQ4	Execution and querying can be quantified and the information systems performance can be measured.	Semantic is the study of meaning of language which is by nature a social construct. As such the field of semantic web technologies is embedded in the interpretivist philosophy.
RQ5		The question refers to social constructs that only exist through the lent of human mind such as organisations, communities, authoritative institutions, services, information and how they relate to each other within the smart city paradigm.

3.3 RESEARCH APPROACHES AND STRATEGIES

3.3.1 Strategies

The research undertaken is part of an active approach where concrete actions are taken to seek answers regarding the research questions. This type of approach is characteristic of the action research, sometimes called applied research, where immediate solutions are given to specific problems [491]. The action research is designed as an iterative process that starts with planning how to conduct the actions or observation, realizing them, reflecting on the outcomes of those and revising the original plan against the first results. This iterative process has been conceptualized by Kemmis and McTaggart in is presented in Figure 3-2 [494].

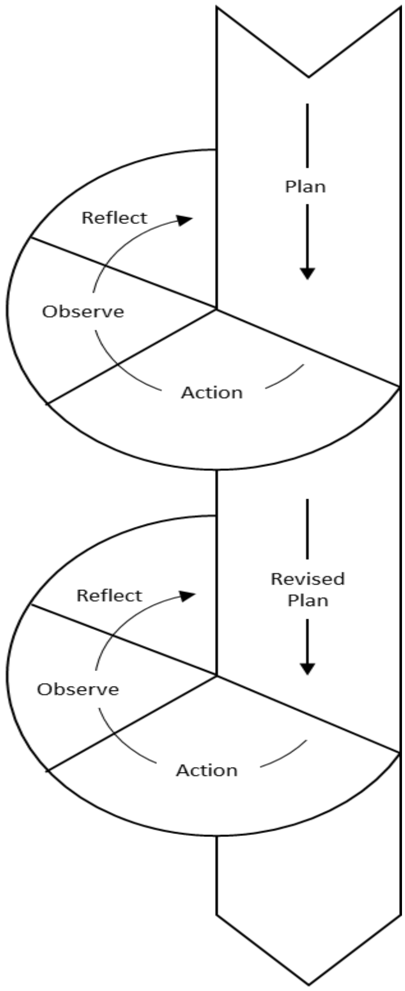


Figure 3-2 Kemmis and McTaggart action research spiral

This approach is opposed to the fundamental research where practical work is not a necessity. The fundamental research is driven by the desire to extend the fundamental basis of a body of knowledge by generalizing on concept. This type of approach is not goal oriented but rather "*gathering knowledge for the sake of knowledge*".

The nature of sustainability and decision making addressed in this research encompasses both social actors and natural objects, conceptual and empirical aspects. One must engage in a participatory approach, involving stakeholders in order to define what does sustainability and how can we improve it and consider the technical characteristics inherited from ICT systems and computer sciences. Therefore, this call for a participatory action research strategy, the most typically used form of research in information systems [495]. The participatory action research is a form of action research characterized by the "*involvement of the practitioners as both subjects and co-researchers*" [495]. This participatory approach is done via the direct interaction with the stakeholder and their inclusion into the loop, the collaboration with experts as well as the conduct of iterative surveys.

A survey itself is another type of strategy that aims at collecting information via the responses to questionnaires or interviews. A survey, following its structural extent, can be either quantitative or qualitative, positivist or interpretivist. In the frame of this research, a survey could potentially help at gathering information from experts on the definition of key performance indicators toward sustainability. In the present case, an intensive and quantitative approach is favoured since the purpose of the survey is to construct a framework that does consensus within the expert community. However, the complexity of the topic of sustainability will require some alternative qualitative questions that aim at better catch a particular perspective.

As well as survey and action research, case studies can serve the research purpose. Myers and Avison define case studies as follow [496]:

> *"A case study examines a phenomenon in its natural setting, employing multiple methods of data collection to gather information from one or a few entities (people, groups or organizations). The boundaries of the phenomenon are not clearly evident at the outset of the research and no experimental control or manipulation is used"*

Case studies have the advantage of contextualising the subject of the research by introducing real-life circumstances which ultimately allows a greater depth of understanding. A drawback is the lack of control over the phenomenon and variables. Equally, cases studies are also strongly related to the design research, a research approach promoted in IS (Information systems) for the creation of business-oriented information technologies innovation [497]. Consequently, case studies are considered in this research to test and validate what has been previously theorised and designed. Some insightful observations will be taken from the case study that will enable to revise the model to effectively answer the research questions. In this perspective, because the case study has a validator status, the sites and structures constituting the case study must be chosen with great care. In the present case, the sites must comply with the intensive need of real-time captured data at the neighbourhood scale to support the calculation of the urban sustainability key performance indicators.

Finally, the intensive literature review presented in Chapter 2 is itself a research strategy. It allows the development of a holistic vision of the field, forming the theoretical foundations. A significant amount of knowledge materials must be critically reviewed and compared, hence the interpretivist stance. In the present case, systematic reviews will be conducted in order to draw pattern in the use of forecasting models or the definition of sustainability.

Moreover, considering the integrative and design-oriented aspect of the solution sought, a review of the existing technologies is much needed as the research less aims at creating new ways than improving the available ones.

Table 3-3 summarises the different research strategies addressed in this research.

Table 3-3 Research strategies checklist

Positivist		Interpretivist	
Experiments		Review	✓
Structured surveys	✓	Unstructured surveys/interviews	✓
Case studies	✓	Case studies	✓
Grounded theory		Action research	✓
Ethnography		Role/ Game playing	

3.3.2 Approaches

The research approaches are related to the place of theory within the research. Saunders et al. [487] give three different types of research approaches:

- *The deductive approach* where theory and hypothesis are first developed before being subject to tests in order to validate or reject them. This is the approach often chosen in scientific research. Therefore, tests being the base for validation, it requires rigour, control and a structured methodology.
- *The inductive approach* where observations are made first, leading to the theory's development. The approach emerged with the social sciences that seek cause-effect links without necessarily having a clear understanding of the social world. Such approach requires high data intensity in order to draw valuable, unbiased theory.
- *The mixed approach* that is to the deductive and inductive approach what the pragmatism is to the positivist and interpretivist philosophy. It is the result of a new perspective that advocates for the use of both inductive and deductive approaches to leverage their respective qualities [498].

The current research project would benefit from a mixed approach. Indeed, the mixed approach has many advantages formulated by Migiro and Magangi [499] that would help in answering questions on sustainability and information systems, two complex domains that covers both scientific and social sciences, positivism and interpretivism, quantitative and qualitative information: (1) the multiplicity of methods, both quantitative and qualitative, can converge and/or combine to answer the research questions; (2) it can response to a greater set of questions concerning the domains; (3) it can lead to the production of a more complete knowledge of complex fields; (4) it frees from constraints inherited by a single method.

3.4 RESEARCH DESIGN

3.4.1 Background

Research is defined as

> "*a quest for knowledge through diligent search or investigation or experimentation aimed at the discovery and interpretation of new knowledge*"[500]

This "new knowledge" is obtained by the adoption of systematic procedures and techniques which ensure a rigorous and reliable investigation and hypothesises validation. The first step in designing a research methodology is to define the purpose of the research. Research purposes are classified into three distinct categories [487]:

- *Exploratory*: aims to explore the different aspects of the research domains in order to gain in understanding of a problem. In this approach the nature of the problem might not be clearly outlined and can be subject to changes, gaining in focus as the knowledge is getting deepen. Therefore, it has the advantage to be flexible and to follow new directions.
- *Descriptive*: aims to describe accurately real-world phenomenon about persons, events or situations. It is greatly used in medical research for the description of a disease for instance [501] or in social sciences where one has to describe characteristics or behaviour of a sample population [491]. However, Saunders et al. argue that descriptive research must be an extension of the exploratory or explanatory research since a description should be a mean to an end rather than an end itself [487].
- *Explanatory:* aims at drawing relationships between specific variables in order to better understand influences. Such researches often rely on statistical analysis to correlate variables. Explanatory research can come after descriptive research in order to catch better insight of a phenomenon. In that case, the studies are known as a descripto-explanatory study.

The current research presents characteristics of the exploratory research as it aims at drawing a new urban sustainability assessment framework and at deepening knowledge on the IoT as support for sustainability assessment, an emerging paradigm.

3.4.2 Design Overview

Figure 3-3 presents an overview of the research design employed in this research. It is adapted from the "Three Cycle View of Design Science Research" developed by Hevner in the same name's article [502].

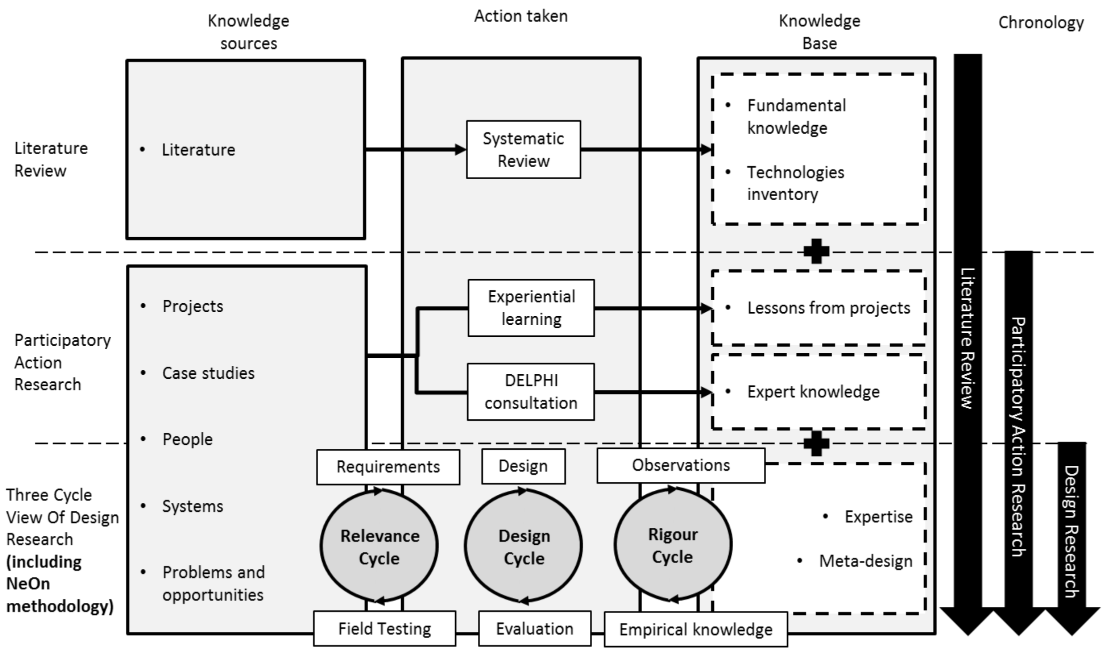


Figure 3-3 Research Design

The design considered extends the design research cycles by integrating the multiple strategies addressed in Section 3.3.1. It is divided into three different methods. The first method consists in an exploratory literature review that aims to reinforce the knowledge foundations. It will help in discovering the research boundaries as well as the eventual gaps that need to be filled. The literature review is here used to initiate the research but is not an end in itself and the second step should be starting relatively quickly, as the action research strategy suggests it. Note that the literature review should not be constrained to this first step and that knowledge materials should be review throughout the research when new elements come in. The second method is typical of the participatory action research strategy where projects partners and organisations are actively involved in the knowledge base development. The knowledge can be gained via structured quantitative online surveys submitted to the field experts for advisory purposes. Additionally, the participation and intervention in the different projects will help in gathering qualitative information through experience, informal exchanges, discussions etc. In the same way than for the literature review, this method is not constrained in time and can be done throughout the entire project once initiated. Finally, the 3rd method relates to the design research strategy where an experimental information system is designed in an iterative process taking advantage of the case studies and incremental gain of knowledge. This last step aims at developing a proof of concept that hopefully will help in fulfilling the gap left in the investigated fields and answer the research questions.

Each strategy is valuable to answer the research questions formulated in section 1.3. Figure 3-4 describes briefly how the strategies implemented address the research questions. More details on how the research questions are addressed are present in the following sections.

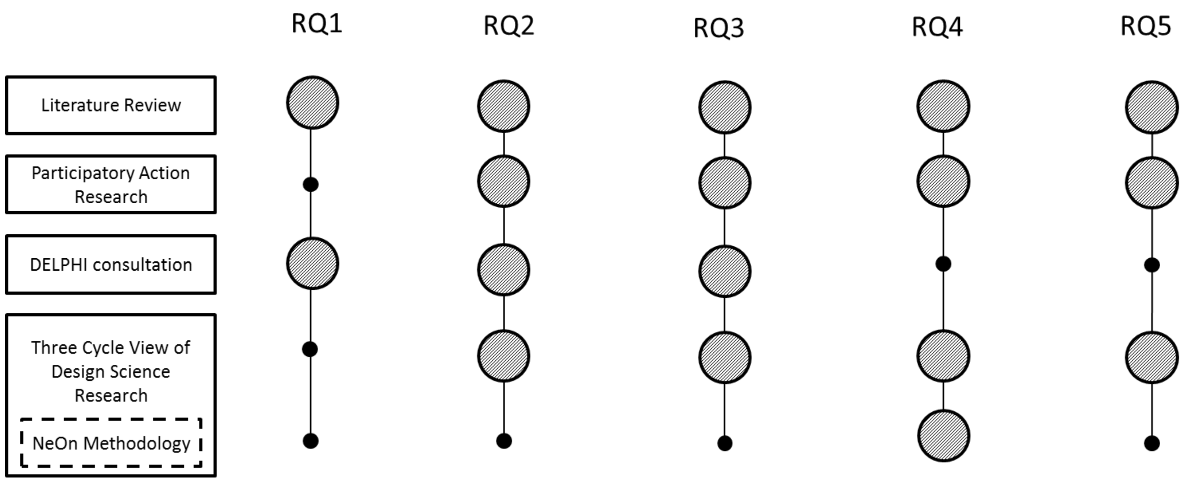


Figure 3-4 Research strategies against research questions

3.4.3 Fundamental knowledge acquisition

The literature review has been conducted for the construction of the knowledge base foundation. It is often the first step in the research methodology since it enables to clarify the boundaries of the research as well as define the eventual gaps present in the field. The evidences collected throughout the literature will help at answering questions in "what" such as "what defines the field?", "what are the actors, techniques, organisations involved in the field?", "what challenges is the field facing?", "what are the current leads to overcome those challenges?" etc. The literature review follows a "natural" progression from theoretical aspects to technical specificities, starting from the definition of sustainability and smart cities to the discovery of appropriate technologies. If the theoretical aspects are investigated at the very beginning of the research, the technologies are introduced incrementally following new directions that the considered solutions take. The literature review will then covers the themes of urban sustainability assessment, internet of thing, smart cities, semantic technologies, smart city value chain and the value of forecasting in decision-making systems.

In order to have a broad view of sustainability and forecasting models, those two domains have been subject to a systematic literature review. Systematic reviews differ from classical reviews with the adoption of an extensive literature investigation and the meta-analysis of it which ultimately reduce biases of misleading studies [503]. In systematic reviews, a precise protocol must be followed in order to ensure clarity, rigour and reproducibility. The authors must, on the base of the research questions, define search criteria to select accurate publications. Once the selection done, the author must first analyse the publications' meta-information and then proceed to a cross-comparison of the various studies' findings [504], [505]. Applied to the sustainability domain, urban sustainability assessment frameworks will systematically be reviewed and the different themes, sub-themes, criteria and indicators, as well as the weighing systems, will be studied with good care in order to draw a new framework where recurrent features converge. Applied to forecasting models, a systematic review of case studies on electrical load forecasting models will be complete in order to find patterns in their uses by the expert community.

The literature review strategy addresses the following research questions by:

RQ1. Drawing conclusions on several gaps and issues currently available in the field.

RQ2. Investigating possible solutions that could help practitioners in their decision making. This includes conceptual re-think of the assessment as well as technological assets such as the smart city, decision support systems and predictive models.

RQ3. Permitting to exhaustively list technologies that could help in capturing in real-time USA KPI. Some researches are presented to demonstrate the feasibility and value of such solutions.

RQ4. Spotting certain issues inherited from the IoT and the smart city paradigm. Especially, the issue of interoperability of information system that is a key component for the realisation of a real-time USA framework. The literature review has therefore brought strong evidences that semantic web technologies are valuable means to tackle the issue.

RQ5. Theorising on this matter with studies on the future of smart cities, their key aspects and actors, and collaborative networks theory.

3.4.4 Knowledge Acquisition from Projects

Engaging in research projects closely related to the original research can be a valuable source of knowledge and can provide a privileged setting to pursuit case studies. The participatory research and the underpinning benefits of such approach have been studied in great details by Kolb in his book "*Experiential learning: Experience as the source of learning and development*" [506]. Klob defined experimental learning as:

> *"the process whereby knowledge is created through the transformation of experience"*

He has theorised the experiential learning process as a four stages cycle, presented in Figure 3-5, including the concrete experience, the reflective observation, the abstract conceptualisation and the active experimentation.

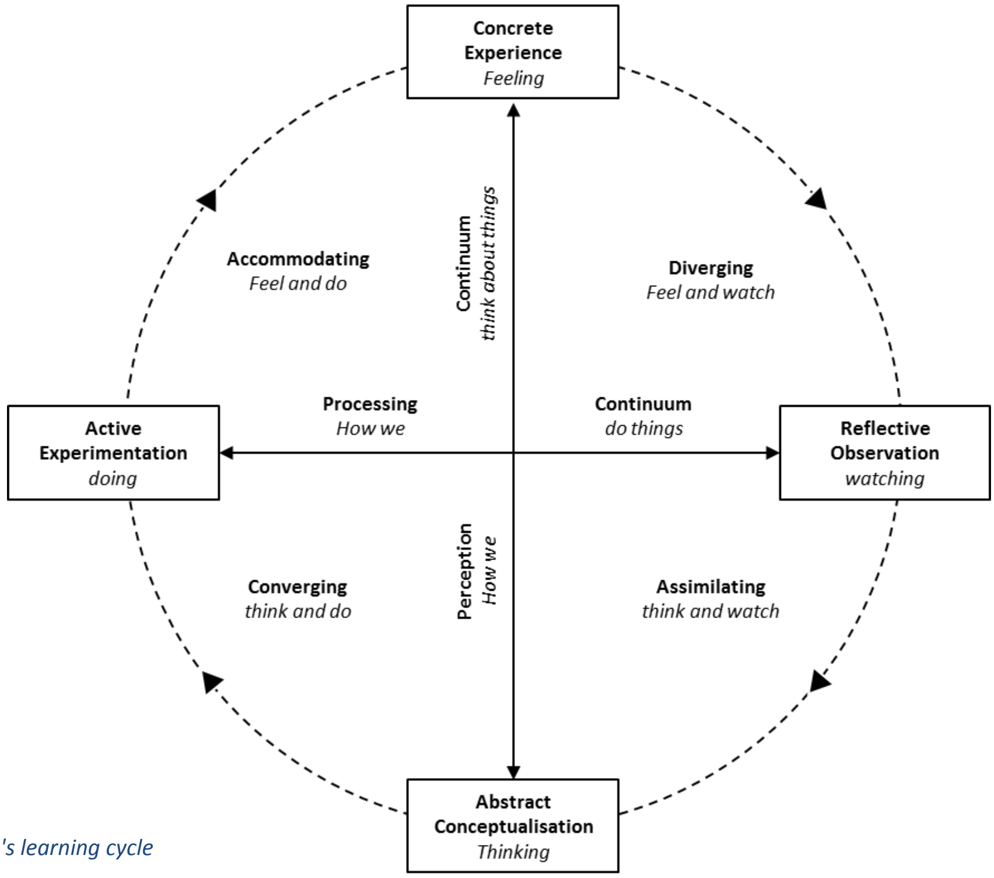


Figure 3-5 Kolb's learning cycle

The concrete experience and the abstract conceptualisation are two opposite means for grasping experience, either by apprehension for the former or by comprehension for the latter. In an analogue way, reflective observation and active experimentation are two opposite way to transform the experience by internal reflexion (intention) or by external experimentation (extension). Combined, the different forms of perception and processing define 4 different types of knowledge, namely, the divergent knowledge, the assimilative knowledge, the convergent knowledge and the accommodative knowledge. According to Kolb, those four types of knowledge constitute the building blocks of a higher form of knowing.

In the present case, experiential learning via projects' participation would enable to grasp domains underpinning knowledge by sensing the underlying aspects related to the actors, goals, structures, methods etc. The different projects have been selected with good care, looking at the domains the cover in accordance with the original research. The prerequisites for their selection were the study of a domain constituent of the sustainability concept at the neighbourhood to district scale, and the integration of intensive and effective ICTs within the cases studied for a real-time assessment.

The Participatory action research strategy addresses the following research questions by:

RQ2. Integrating projects that worked on smart solutions for energy delivery which enable to picture what is being done in the field in term of real-time decision support e.g. optimisation, prediction, user friendly environment etc.

RQ3. Being information intensive and showcasing how ICTs can support assessments and decision making. Hence helping investigating the relevance of the smart city paradigm in the field.

RQ4. Addressing the matter of interoperability along with the inclusion of semantic web technologies. Integrating project therefore enables to describe how such technologies should be considered.

RQ5. Enabling to observe the collaborative vision of most projects and how diverse organisations can come together for the delivery of services to stakeholders. This gives insights, to some extent, on how future services can be designed, implemented and delivered in the up-coming smart city paradigm.

3.4.4.1 *Thermoss*

THERMOSS [507] is an EU funded research project that aims at promoting outstanding solutions for the deployment of district heating and cooling (DHC) technologies in Europe. The project is part of a commitment of the EU authorities to enhance the energy efficiency of the residential building stock and to reduce GHG emissions. The project gathers internationally distributed organisations from industry and academy that bring their knowledge expertise to contribute to valuable solutions. The objectives are multiples:

- Ensure that heating and cooling supply match the district demand
- Improve residential building thermal energy systems
- Raise construction and energy awareness on open data for data-driven solutions
- Promote reproducible solutions throughout Europe for an effective market integration

In the perspective to fulfil those objectives, the THERMOSS project proposes a smart DHC approach with a twofold contribution, an innovative DHC technologies implementation along with two decision support tools for a successful integration and optimisation. A 5 steps development process have been designed (Figure 3-6) starting by the creation of technologies' database.

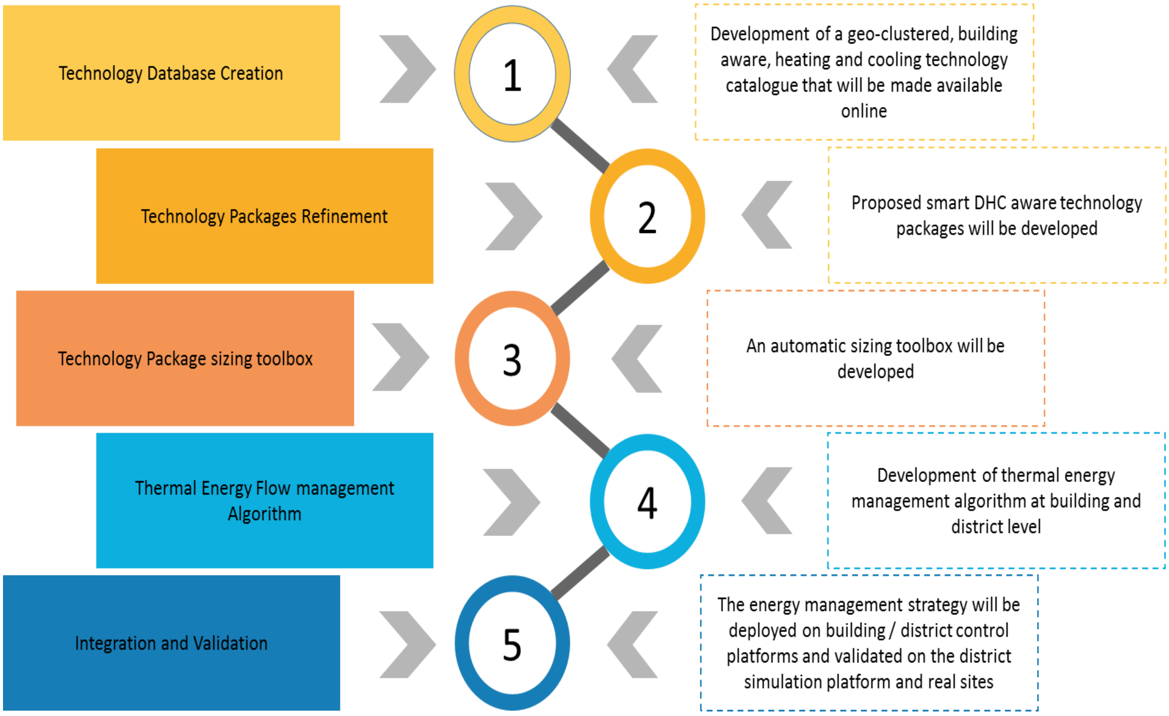


Figure 3-6 THERMOSS development process

- (1) This geo-based catalogue would enable the selection of appropriate DHC technologies available in particular settings. A set of 16 parameters covering climate, building typologies and socio-economic aspects have been defined for the creation of up to 116 geo-clusters that will ease the DHC technologies collection.
- (2) A subset of DHC technologies will then be selected according to step 1 for the pilot sites implementation. Concretely, four technology packages are considered: micro-CHP (peak load boiler integrated), gas absorption heat pump, hybrid heat pump with gas boilers as a backup and air sourced electric heat pump. They will be

combined with additional technologies such as thermal solar panels, thermal storages (water tank & building inertia) and two-way substations.

- (3) The third step will focus on the development of a sizing toolbox for the technology packages. An optimisation algorithm will consider local regulations and technologies characteristics such as dimensions, power production and control settings, and propose a cost-energy-comfort optimal solution for the technologies integration. The aim is to help stakeholders of the various pilot sites in the selection and sizing of the technology packages. Such approach, if applied to particular technology package and pilot sites, has been designed to be reproducible to potentially any kind of DHC technology and any European location.
- (4) In a fourth step, a thermal energy management algorithm will be created at the building and district level. At the building level, the algorithm will optimise the thermal energy usage by efficiently integrating thermal storage strategy within the systems. This will require the development of prediction models of the building heating and cooling demand and the energy units' production for predictive control. At the district level, the algorithm will consider the district as a whole and try to minimise energy production by fully introducing the distribution system as storage. This approach will enable thermal peak shaving relying on storage. Predicted schedules will then result and be given to the plant manager in order to adjust the production.
- (5) Finally, an implementation of the above strategies will be done with special care on the integration of ICTs within the system. Indeed, ICTs are at the core of the system to transmit reliable information from the production units and buildings to the sizing toolbox and optimizers. In this prospect, attention must be put in communication protocols and standards, as well as unified information systems that homogenise information, flows between heterogeneous organisations and applications.

3.4.4.2 *Pentagon*

PENTAGON [508] is another EU funded project that aims at promoting innovative energy conversion systems between thermal, gas and electricity, drawing a new generation of eco-district. Such approach is introduced to answer the current challenge of heterogeneous renewable energy sources, bringing better flexibility to the electrical systems. Various international organisations collaborate to fulfil a set of objectives:

- Meet new requirements of eco-district energy systems by introducing Power-to-Gas technologies.
- Leverage the flexibility capacities of individual buildings via the development of multi-objective energy management program that synchronises demand, production, storage and conversion.
- Showcase multi-scale ICTs as a mean to support live assessments and smart energy management.

Similarly to THERMOSS, PENTAGON axes its contributions around two mains inputs that are the promotion and implementation of cutting-edge technologies such as power-to-gas and power-to-heat technologies coupled with optimisation platform and simulation environment for an efficient energy management.

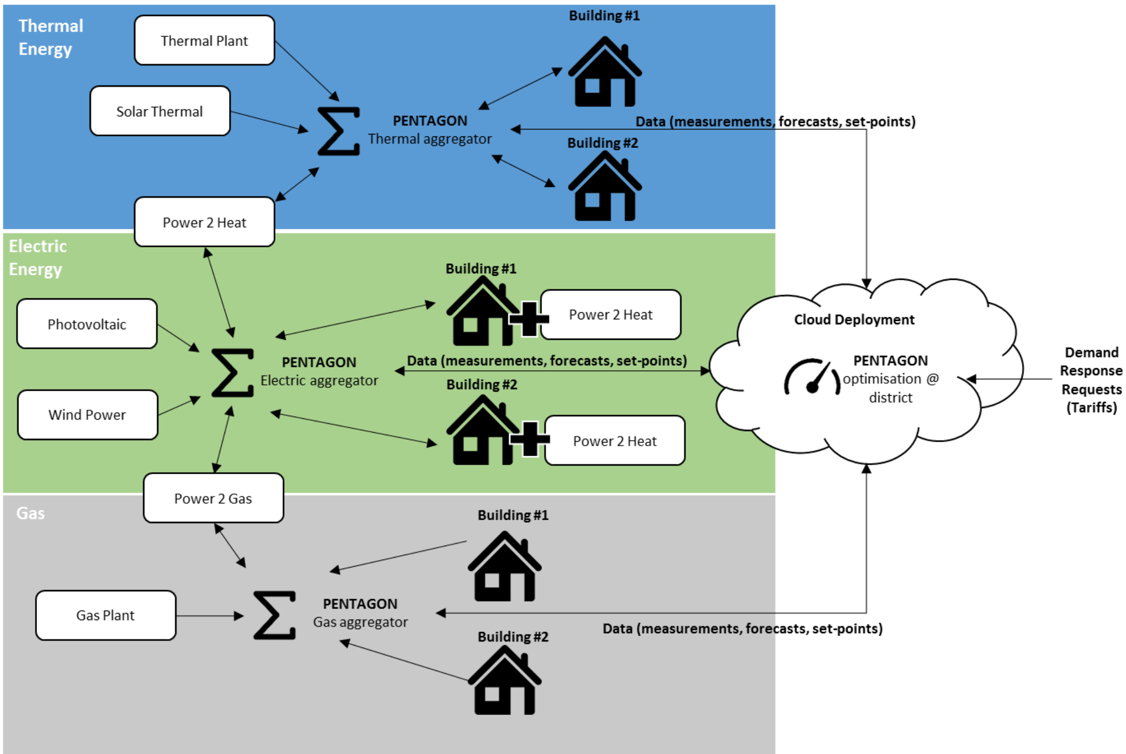


Figure 3-7 PENTAGON eco-district concept

As Figure 3-7 presents it, the project takes into account three heterogeneous energy sources that are the electricity sources with PV panels and wind power, the heat sources with solar thermal and thermal plants and the gas sources with a gas plant. The innovation comes from the power-to-gas and power-to-heat technologies that bridge those different types of energy carriers by converting one into another. Therefore, every production units are potential energy sources indifferently of its type. Electricity excess can be stored either by buildings thanks to their thermal storage capacities or by gases. This confers more flexibility to the electricity grid. However, such approach can only be fully efficient when supported by a performant management system that precisely decides when a conversion is needed and to which extent. For that reason, the PENTAGON project equally designs some optimisation algorithms that take as input the aggregated energy consumption and production as well as previously calculated forecasts and it outputs control settings for energy production and conversion/storage.

3.4.4.3 *CUSP*

The Cardiff Urban Sustainability Platform (CUSP) is an on-going project of the BRE Trust Centre for Sustainable Engineering at Cardiff University. CUSP is a decision support tool that provides insightful urban analytics in an immersive interface. The 3D web-interface (Figure 3-8) of the site of Ebbw Vale enables to monitor in real-time energy consumption and production to support energy planning and flexibility.

Figure 3-8 CUSP interface on Ebbw Vale case study

The IFC format has been chosen as a privileged standard for building data exchange in the platform. 3D IFC models are produced from a 3D point cloud collected by the mean of laser scans. The scans are imported into Autodesk Revit and used as an underlying layer for the drawing of 3D BIM models like shown in Figure 3-9. The metadata hold by IFC models are then the base for a semantic-based approach.

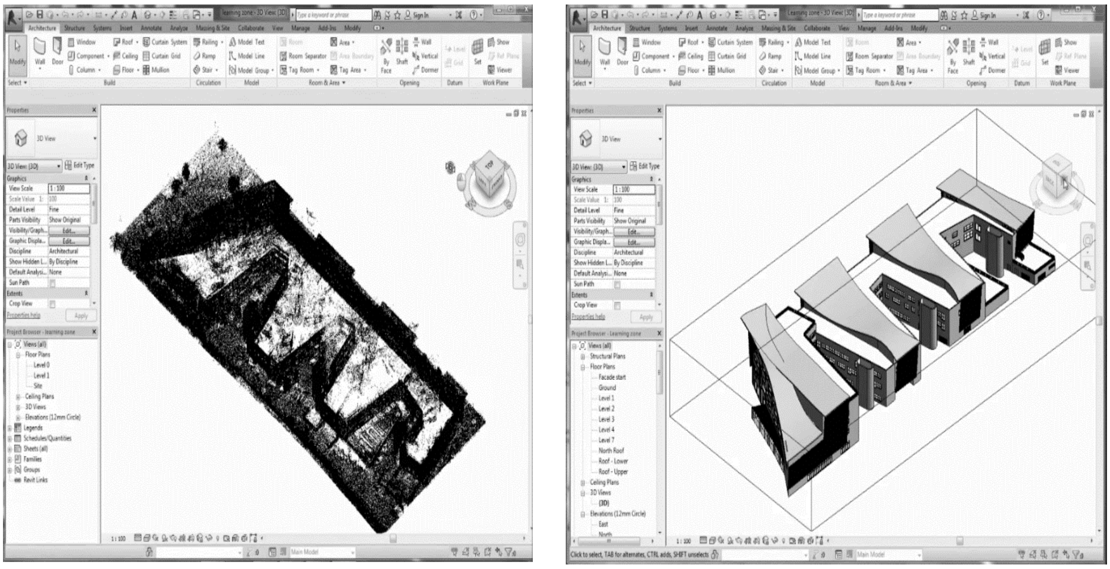


Figure 3-9 IFC model production using Autodesk Revit

CUSP is built using the Unity Game Engine due to its cross-platform deployment with efficient use of computational resources. 3D building information models are then exported into Unity, on a 3D terrain which reflects the actual topography of the area enhanced with high-resolution aerial photography. These 3D models are enriched with semantic information across domains and scales.

Some dashboards provide the direct outcomes of the energy and water domains by clicking on a specific building (see Figure 3-10). They can share real-time information as well as a day-head predicted demands made by machine learning algorithms. In the energy dashboard, the operational schedule for each energy production unit, key energy performance indicators, are displayed. In the water dashboard, alerts to warn of existing or predicted issues such as water quality, flood risk or sewer flooding and network leakages are displayed. This provides useful insight into the management of the network and allows the manager to make an informed decision of the day ahead.

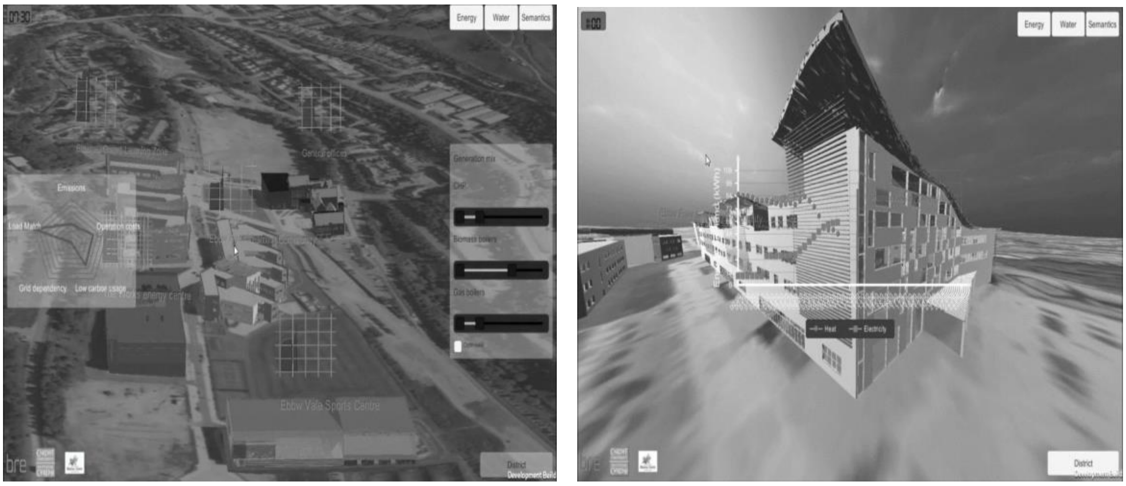


Figure 3-10 Energy-related dashboards

3.4.4.4 *Mas2tering*

MAS2TERING [509] is an EU funded research project that aims at promoting cutting-edge ICT platform supporting monitoring and optimisation of smart grids low-voltage distribution. Besides the technological aspect, the project ambitions with its platform to create a new business model between grid operators, telecommunication and energy providers companies. Mas2tering stands for Multi-Agent Systems and Secured coupling of Telecom and Energy gRIds for Next Generation smart grid services) and is built onto five key pillars presented in Figure 3-11. The first pillar "*Interoperability and standards"* focuses on the ability to seamlessly exchange information no matter the communication protocols and standards used. This is essential in decentralised smart grids since it enables heterogeneous systems to communicate efficiently. For that, the project has considered the implementation of a Telecom smart gateway that can collect information from smart appliance as well as receive some request for control long with standardised interoperability mechanisms such as the FIPA-ACL [510], the IEC 61968/61970/62325 [511] or the OpenADR (Automated Demand Response) standard [512]. The 2nd pillar concerns the software component "*Reliability and security*". Indeed, as smart grids grow in size, cybersecurity has become a key aspect to ensure reliability and quality of supply. MAS2TERING has taken a technology-driven approach that aims to ensure secure authentication and communication between heterogeneous environments. Consequently, risks of breach of privacy and security from insecure authentication have been investigated as well as data losses from ICT and their influence on the grid reliability and QoS. The outcome is the conception of guidelines for smart grid security management part of an effort for standardisation. The 3rd pillar "*self-organising architecture"* refers to the holonic perspective of smart grids considered in the project where data and power flow can be bi-directional. The grid is composed of autonomous systems communicating one with another to pursue an effective flexibility, scalability, resilience and openness. The design of holonic

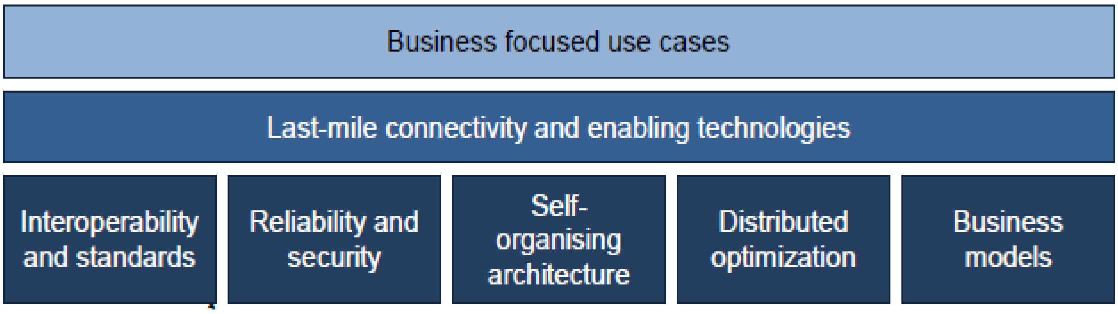


Figure 3-11 MAS2TERING key components

architecture with the support of multi-agent systems for the network optimisation has been studied and validated through several case studies. In the 4th pillar "*Distributed optimisation*", the project has developed decision support algorithms that help to balance energy distribution balance between production, storage and consumption. The objective is to ensure a cost-effective and resilient power grid by avoiding peak loads, reducing energy losses and use local, preferably renewable electricity sources. Finally, the 5th pillar of the project is more business oriented. Indeed, MAS2ERING engaged efforts in developing a new business model for LV smart grid, redefining the role of stakeholders, energy and data services. In this new approach, special care is given to the role of prosumers and the mean to aggregate their inputs as well as the key role of Distribution System Operators (DSOs) and their ability to bring flexibility to the grid.

3.4.4.5 *Lessons from projects*

All the previously cited projects have been a valuable source of knowledge essential in the pursuit of the original research. Some of the tasks the author has been involved in present interests directly or indirectly related to the research. Table 3-4 presents the author contribution to the different projects and which key aspects those contribution covers for the original research.

Table 3-4 Research project inputs

Key research aspects	THERMOSS	PENTAGON	CUSP	MAS2TERING
Urban analytics for sustainability assessment		- Consultation on urban sustainability KPIs		
Urban energy systems		- Review of energy-related KPIs	-Building performance simulation using Design builder and Energy+	
Web semantic technologies	-Development of a semantic repository for data access (Ontology-based Data Access (OBDA))			
BIM models	-3D BIM IFC models created from floor plans with Autodesk Revit		-3D building scan within laser scanners -3D BIM IFC models created from scans with Autodesk Revit	
GIS models	-Creation of geo-cluster for DHC technologies			

	selection based on building typologies, environmental, social and economic criteria.			
Prediction models	-Development of prediction models for heating and cooling load at the district level		-Development of prediction models for energy load at the district level	-A systematic review of forecasting models for electricity load prediction
Decision support tool			-Design of a decision support tool	

Beside those clearly defined contributions, informal knowledge has been acquired from the projects. Having the possibility to exchange with domains experts, to face technical and organisational issues has participated to better picture certain limitations and challenges that face the domain. Additionally, it allows to have a critical view of the current state of the field by catching with aspects are left aside and which ones are emphasised.

3.4.5 The DELPHI Consultation

At the stage of the research, the author must provide evidences to answer the question: *How an effective urban sustainability assessment can help the different parties of a city in their decision making?* This research question is addressed via the use of a DELPHI survey. The DELPHI technique is particularly relevant in the case of an exploratory research where complex issues involving new or future tendencies are present [513].

The DELPHI method first appeared in the American business community and is nowadays a well recognize methodology in several key sectors such as health, defence, education, transportation, engineering etc [514]. It is an effective systematic method that allows the collection of data from experts in the field of interest. Its effectiveness lies in its features:

- *Anonymity* of the respondents preventing form bias and external influences
- *Iterative* process allowing the definition of a consensus toward the question
- *Controlled feedback* communicating results between rounds
- *Statistical group response* that assures the equity between members in the final response.[515]–[517]

Three types of DELPHI studies can be identified [517]:

- The "classic DELPHI" where data are collected from experts in iterative rounds with feedback between each round until consensus is achieved.

- The policy DELPHI which is mainly used in social sciences for social or political questions. The outcome is the creation of policies through structured public dialogue.
- The decision DELPHI is embedded in a decision-making process. It aims to coordinate thoughts in order to come out with a definitive decision. This type of DELPHI study includes five distinct features: quasi-anonymity (the experts are known by name to everybody from the beginning of the study); iteration; feedback; statistical group response; and constancy in responses among the experts on a specific issue [518].

For the current research, a "classic DELPHI" is considered. Figure 3-12 presents the different stages of the expert consultation undertaken. A near-real-time sustainability assessment framework is first developed based on the systematic review of urban sustainability assessment schemes present in the literature. Then a pilot version of the survey is draft using the online survey tool Bristol Online Survey [519], a recognized tool for the creation and design of the online survey. The pilot survey is distributed to a small panel of expert that will give insightful feedbacks on the goodness of the survey in term of design, length and understanding. Subsequently, this will result in a meta-analysis of the survey to determine if it is well designed or not. Such method, taken in an iterative approach will end with the development of a consistent and user-friendly online questionnaire. Once the final survey designed, it is distributed to a new and larger set of experts. Several rounds will be done in order to bring consensus, fill certain remaining knowledge gaps and deepen some fuzzy aspects on the development of a near-real-time sustainability assessment framework. After several rounds, the framework definition should be refined and agreed by experts and therefore validated.

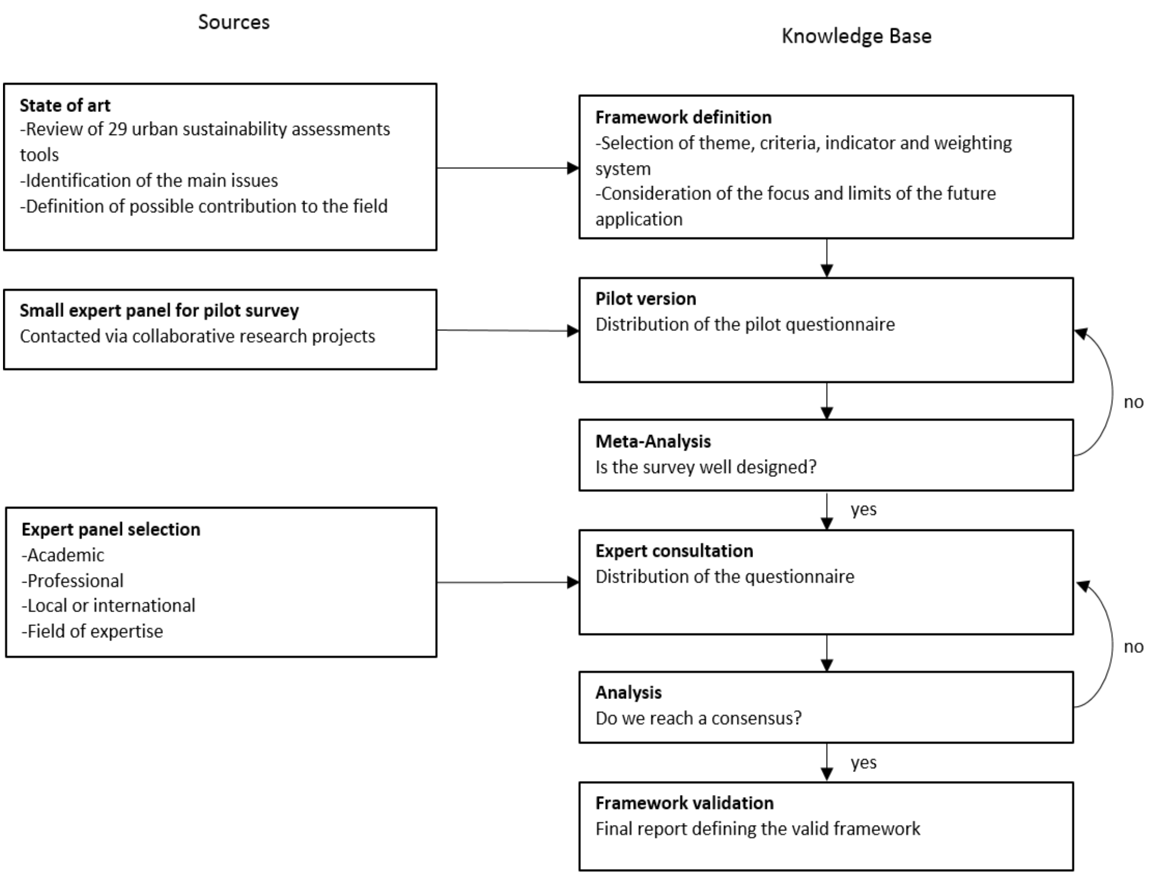


Figure 3-12 Delphi methodology

The survey strategy addresses the following research questions by:

RQ1. Allowing experts to give their opinion on the sustainability KPIs and on the main issues when evaluating sustainability at urban scale.

RQ2. Providing open questions that allow experts to open on best practices in the fields, giving valuable information on an efficient assessment.

RQ3. Exploring the feasibility and cost to capture KPIs in real time using ICTs.

3.4.6 NeOn methodology for ontology development

At this stage, one must answer the research question RQ4 "*How can semantic web technologies unify heterogeneous data resources for holistic services and applications?".* Semantic web technologies being considered to solve this issue, the NeOn methodology for ontology development demonstrates how such technology can "unify" heterogeneous information sources by integrating reusable knowledge resources within ontology design process and aligning them.

An ontology being by nature an integrative information system, that considers other already existing ontologies, a specific methodology must be followed in order to efficiently develop an interconnected knowledge network. Ontology development methodological aspects have already been the subject of many studies (see section 2.3.4), this section will therefore give the outline of the selected one, the NeOn methodology.

An often cited methodology for ontology development is the METHONTOLOGY [256], [368], [520]–[522] which confer it a reference status in the field. However, a study undertaken during the NeOn methodology conception [379] has highlighted was considered superior to its parent the METHONTOLOGY because it benefits from (a) ease of understanding, (b) scenario-based approach, and (c) availability of supporting documentation [379]. For those reasons, the NeOn methodology has been chosen to develop the urban sustainability assessment ontology.

The NeOn project primly aims at providing a complete framework for ontological development [378]. The main outcomes of the project are [380]:

- The NeOn glossary which identifies and defines processes and activities for ontology development.
- Nine scenarios of ontological development.
- Two life cycle models that summarize ontology development processes and activities.
- A set of methodological guidelines [381].

Figure 3-13 presents an overview of the NeOn methodology life cycle here solid black components are mandatory steps and dotted ones optional, depending on the scenario. This diagram demonstrates the iterative nature of the ontology development where the designed model must incrementally be improved and re-engineered.

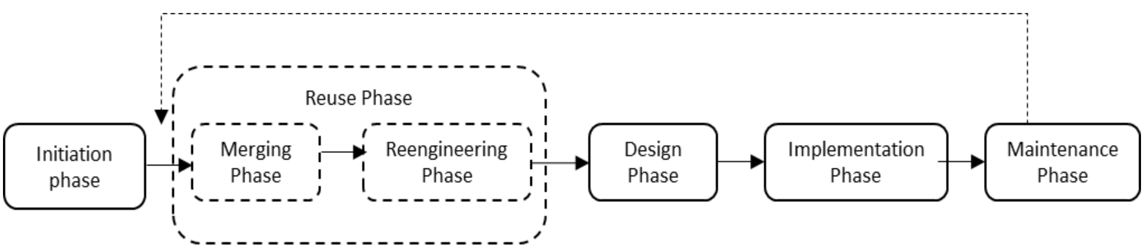


Figure 3-13 NeON methodology [379]

3.4.6.1 *Ontology requirements specification*

The initial phase is probably the most important step in ontology development. It must be considered with good care since conception mistakes at this stage could lead to a poor conceptual and structural design of the ontology. In order to avoid such misconception, the NeOn methodology promotes the use of the Ontology Requirement Specification (ORS), a concept native of the METHONTOLOGY [381], [523]. The ORS objective is to bring clarity to the domain requirements and aspect that the ontology must cover. It is composed of 3 distinct steps: (1) Identification of the purpose of the future ontology; (2) Identification of the intended uses and users of the ontology; (3) Identification of the set of requirements that the ontology should satisfy [524]. Additionally, it is important to define the scope of the ontology and the domain addressed.

The definition of the requirements can be made efficiently by the use of competency questions. Competency questions have been already greatly used for ontological development [373], [381], [521], [524], [525]. They are simply questions that the ontology should be able to answer when queried. To begin, the questions are defined in an informal way with a natural language. This will help in drawing the scope of the ontology as well as grasping recurrent terminologies. Then, once the author has a better picture of the domain, formal questions are formed where exact terminology, properties, links and axioms are extracted. The question should be organised in a structured way with different levels of abstraction from simple questions to complex ones [525]. In that way, requirements move from a fuzzy and uncertain definition to a more specific and valuable system.

3.4.6.2 *Scenarios*

Figure 3-13 gives a simple overview of the ontology development process. However, in fact, the procedure is more complex and greatly depends on the domain considered and the extent to which existing ontologies are reused. The NeOn project has identified 9 scenarios that one might follow in the creation of an ontology. Figure 3-14 presents all 9 scenarios and their interrelationships identified within the NeOn methodology. The different scenarios mainly diverge on the uses of knowledge resources and especially existing ontologies. Following the extent on which the existing ontologies fulfil the new ontology requirements, the developer has the possibility to simply create direct alignments, to re-engineer it or to reuse the present design patterns. In the case of the 1st scenario, it consists of the core activity of ontological development and therefore must be combined with another scenario. Those scenarios highlight the importance of resources reuse and especially ontology reuse, demonstrating the conceptual interest of the field for the creation of a vast interconnected knowledge base.

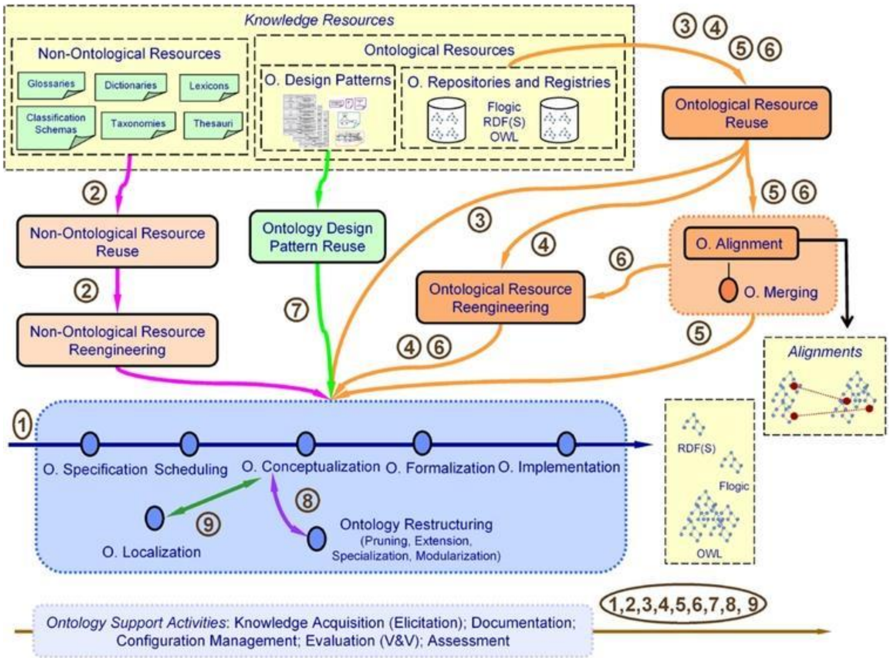


Figure 3-14 NeON scenarii representation [381]

3.4.6.3 *Resources Reuse*

By nature, an ontology is the formal representation of a domain knowledge and consequently is subject to a certain rigour. In this prospect, formal knowledge sources are essential for the development of the ontology. They will define the terminology to be employed and assure correctness. Knowledge resources have been divided into two categories: non-ontological and ontological resources.

The non-ontological is the naturally encountered procedure to extract knowledge about a domain using materials such as glossaries, taxonomies, thesauri etc. They are mostly studied in order to scope the domain, gain terminologies and key concepts useful for the ontology.

Ontology reuse presents the key advantage to be already formalised thus less time and cost consuming during the development stage. Moreover, reusing existing ontologies is in straight line with the ontology field paradigm of an interconnected knowledge base. If the ontology developer is free to use any methods, most (if not all) expert recommend the reuse of ontology when applicable.

Simperl [372] gives the main steps to follow for efficiently reuse ontologies:

- (1) *Discovering ontologies*: a tremendous amount of ontologies already exists about various domains. The search of eligible ontologies is fortunately helped by the use of ontologies search engine such as such as Swoogle [526], Watson [527] and/or by ontology repositories such as the DAML ontology library[528] or the Protégé OWL library [529]. Additionally, some upper-level ontologies presenting high-level concept exist such as SUMO [530], DOLCE [531], BFO [532], CIDOC-CRM [533] and often serve as a base for abstract concepts.

- (2) *Selecting those to be reused:* once a set of ontologies has been chosen, the developer must identify which ones complies with the newly developed one. For that, ontologies can be reuse entirely or partially depending on the aspect they cover, the selection being based on ORS previously defined. Therefore, the ORS and competency questions must be well-defined in order to efficiently select valuable ontologies. Note that broad ontologies tend to over-generalise hence missing some specific domain aspects while really detailed ontologies convey intensive and hardly apprehendable knowledge. Thus, the author must have a good idea of the level of abstraction one wants to give to its model and considers potential costs and benefits of reusing existing ontologies or developing one from scratch.

- (3) *Customization of relevant ontologies*: after the ontologies ‘selection, the developer often has to modify them in order to suit with the intended purposes. To do so, one can simply add or remove certain axioms, restructure the architecture or translate from a language to another for example. Depending of the ORS, the ontologies can be re-engineered or reuse as-is.

- (4) *Integration into an application ontology*: finally, the last step is the alignment of the different domain ontologies to form a new one. The author must map certain concepts together via equivalences, potential restrictions and/or properties. The end-model must then prove consistency and be reworked in an iterative way between the steps (3) and (4) until it reaches soundness.

3.4.7 Design Research & Real Case Studies

Case studies are worthwhile strategies to explore existing theory [487]. It allows a contextualisation of the theory and might provide evidence to confirm or infirm the initial hypothesis. Moreover, case studies enable the researcher to gain informal knowledge

[534]. In the particular context of the research where data are at the core of the system, a case study is a valuable source of exploitable data. Several requirements are to be considered when choosing a case study: a sensor network must be integrated at the district level; historical data must be accessible as well as streams; available data must support the calculation of certain KPIs.

Cases studies are fully integrated in the Three Cycle View of Design Science Research where a system design leverage on real case studies for iterative tests.

The design research strategy along with case studies address the following research questions by:

RQ2. Opening on how real time information are processed as well as the inclusion of predictive models within the tool. Testing it gives insights on the value of certain technologies for decision making.

RQ3. Designing and implementing the services and demonstrating the usefulness of ICT for sustainable assessment.

RQ4. Implementing semantic web technologies within the design and making iterative tests that allow demonstrating their value for interoperable information systems.

RQ5. Bringing organisational issues related to the case studies, for instance data access and privacy. Case studies can therefore be helpful in understanding how human, legal and financial aspects could affect smart services provision.

The following sections present the case studies that have been considered to support the design research strategy.

3.4.7.1 *Ebbw Vale, The Works*

Figure 3-15 Ebbw Vale, The Works site

The site of Ebbw Vale called "The Works" (South East Wales, UK) has been considered for the framework's validation. The site was formally occupied by steelworks which has been closed in 2002. A program of regeneration has then been initiated and it is in 2012/2013 that a new district composed of 6 main buildings emerged (Figure 3-15). The district now hosts the General Offices that included the Gwent Archives and the local council, a Learning Zone campus and a school where middle school, A-level education and vocational courses are provided, a leisure centre with swimming pool and sport facilities, a car park and an energy centre that provide heating and electricity to the district. The site is a state of the art £350 million project that integrates excellence BREEAM certified buildings, efficient and low carbon energy generation systems and a vast sensor network that records energy, climate and comfort related data within the entire neighbourhood. As shown in Figure 3-16, the Energy Centre is ground for a Combined Heat and Power (CHP) plant of 400kW thermal and 375kW electricity, two 500kW wood pellet boilers and four gas boilers of 1750kW each. Additionally, two water tanks are connected to the biomass boilers that serve as buffers for a continuous operation.

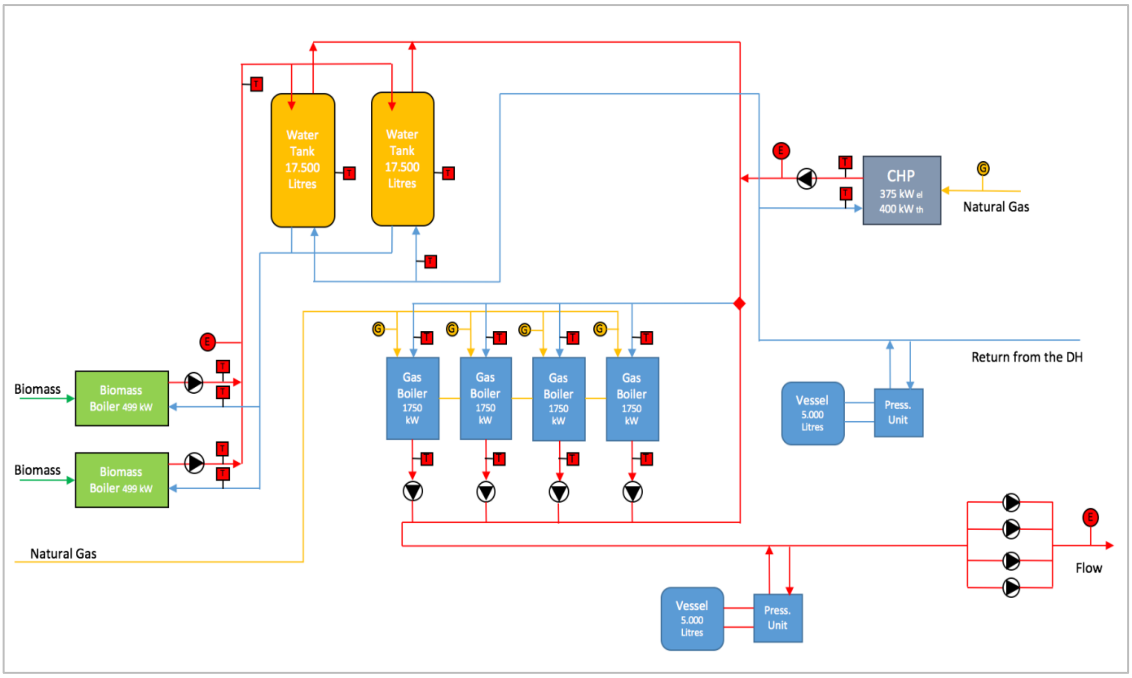


Figure 3-16 The Work Energy Centre schematics [678]

Figure 3-17 presents the district heating and electricity network. Heating is supplied entirely by the Energy Centre that can cover up to 12MW loads over the district. Western Power provides 8 MW electricity that goes along the low voltage network sourced by the Energy Centre. While all buildings but the Car Park are supplied with heating, only the Energy Centre, Learning Zone and the Car Park are connected to the low voltage network. The buildings are connected to the heating network via two heat exchangers, one allocated for space heating and one for domestic hot water. In the case of the Leisure centre, the heat exchanger allocated for space heating is equally allocated to the pools.

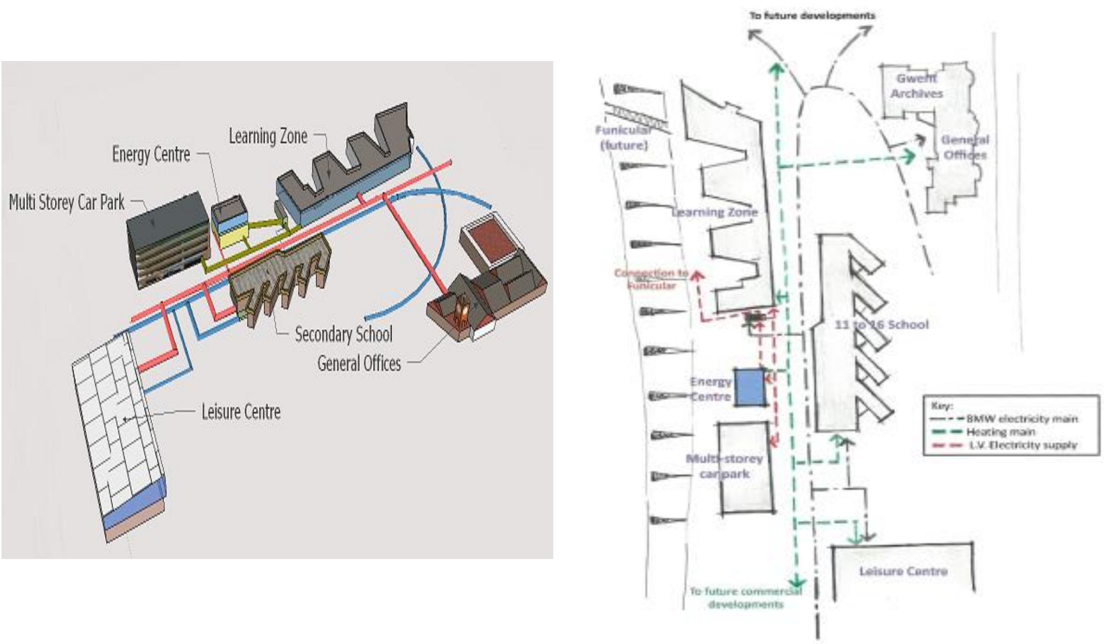


Figure 3-17 Ebbw Vale site District Heating and electricity schematics [678]

The site of Ebbw Vale is particularly interesting in the frame of this research as it monitors highly frequently energy-related data. Indeed, all buildings heating and cooling, as well as electricity loads, are logged 15 min to hourly throughout the entire year. Such measurements are done at system to subsystem levels, allowing a flexible aggregation/desegregation. Same goes for the indoor environmental conditions wherein each room are monitored indoor temperature and humidity levels. A weather station is present on site (on the Coldmills School) looking at air temperature, dew point temperature, barometric pressure, humidity, global radiation, wind speed and direction. Finally, all energy sources units are set with sensors that check heating network flow and return temperatures, gas and biomass consumption, electricity and heat generation. Those features, even if they do not cover the entire field of sustainability, are believed to support a significant amount of KPIs (mostly energy-related) which ultimately can validate the semantic-based near real-time system operation on a subset of KPIs.

Moreover, the choice of such case study is also motivated by its full integration within collaborative research projects where agreements have been defined to freely access the data. This aspect is critical as data are at the core of the research. To be part of a formalised structure where actors are fully aware of the provenance and the destination of data allows the researcher to overcome certain organisational and technical challenges that may occur in another setup. In the present case, if a direct data stream from the sensors seemed to be impossible due to legal and technical limitations, historical data samples were provided.

3.4.7.2 *52 The Parade*

52 The Parade is a Cardiff University-owned building that hosts the BRE Trust Centre for Sustainable Engineering. Within the building are present three open working spaces, a kitchen, a meeting room and a server hub. The building has been refurbished in 2014/2015 and now host energy efficient heating system controlled via BMS. Moreover, the set of sensors have been installed within the building in order to assess energy and water loads as well as indoor environmental conditions. A heating control system has been designed in the perspective to lower energy consumption.

The building has been used in the frame of this research as it provided accessible live stream data. Indeed, despite the fact that the building cannot be equated to a district, this feature is essential to test the validity of the model over real-time data. It is therefore sued more to showcase the model abilities than the assessment performance. Data are accessible via a KairosDB server, a fast scalable time series database [535], REST API deployed by the BRE research team which ensures full access and efficient maintenance of the system.

3.4.7.3 *THERMOSS pilot sites*

Presented in Section 3.4.4.1, THERMOSS is an EU funded research project that promotes cutting edge technologies for district heating and cooling. The project collects data from 4 different pilot sites:

- In San Sebastian (Spain) where the URBEROA district heating from the different gas boilers, CHP and biomass boilers are measured as well as at the substations;
- In Southampton (UK) where a multiple-dwelling building heating provision, individual dwelling heating demand and indoor environment are measured;
- In Portsmouth (UK) with a similar setting of the Southampton site;
- In Chambery (France) where a laboratory setup enables to simulate district heating and measures the different loads, demands and environmental conditions.

All the data are centralised within a “Smart Connector Server” deployed by one of the project partners. The Smart Connector Server goes along a REST API that allows its user to retrieve the data and the sensor metadata either by selecting the last value or by chunks.

3.5 ETHICS

Some ethical considerations are inherited from the field of computer sciences. With the recent growth of data, sensitive data has become a prime consideration. Computer scientist must be trustworthy with special care to maintain the privacy and integrity of data, especially in the case of personal data [536]. In computer sciences, the three main influencers on ethical decisions are (1) the scientist own moral code, (2) informal ethical code in the work environment and (3) formal codes of ethics to be followed [537]. At Cardiff University, the "Research Integrity & Governance Code of Practice" formalises research ethics guidelines that the researchers must follow [538]. Such guidelines along with ethical and data protection training provided ensure the respect of ethic required within the computer sciences field.

Data protection is equally an important consideration when it comes to conduct a survey. Before engaging in a survey, the researcher must get an ethical approval that considers research design, topic, data collection, processing, storage and usage [487]. Cardiff University policy requires that all researches involving personal data must first be subject to ethical review and approval.

3.6 SUMMARY

Mixed approaches have been followed in this research for the construction of the knowledge base and the collection of evidences. An extended literature review of the urban sustainability schemes, forecasting models, decision support tools, smart cities and semantic web technologies has provided a solid background for the contextualisation of the study. The participatory action research has allows the author to engage with experts in the domain and provided valuable case studies that could support the implementation of the real-time urban sustainability assessment framework. Furthermore, a DELPHI consultation has been developed and experts from projects or international have been contacted to validate the relevance of the new framework. Ontological development has required the use of a well-defined methodology, the NeON methodology, as the domain is by nature complex and requires strict procedures. Finally, the "Three Cycle View of Design Science Research" often used in the Information System field has been followed for the iterative design of the USA application, building expertise and requirements through the iterations. Following the urban sustainability assessment schemes review observation, the following chapter will introduce the development of a new scheme as a synthesis of the existing ones. The new scheme will be validated via an expert DELPHI consultation that will question the relevance and feasibility of capturing the addressed indicators in real-time.

Urban Sustainability Assessment Framework for Real-Time Assessment

This chapter will explore the gaps found in the literature concerning current Urban Sustainability Assessment schemes. A new framework will be created and introduced as a mean to reconcile sustainability assessment and smart city.

The following research questions will be addressed here:

RQ1. What are the issues that face the different stakeholders of an urban system toward sustainability assessment?

RQ2. How an effective urban sustainability assessment can help different parties of a city in their decision making?

RQ3. How can sustainability assessment leverage the smart city paradigm, specifically ICTs and the IoT?

4.1 OVERALL DESCRIPTION

The systematic literature review of Urban Sustainability Assessment frameworks presented in Section 2.1 has highlighted some key issues that still need to be tackled. Indeed, despite considerable progress on the key indicators addressed, supposed to define sustainability at the urban scale, there is still no real consensus on what should be considered. If some issues relating to the environmental dimension such as water scarcity or energy savings are well recognised and universally accepted, other aspects more abstract still do not reach such agreement among experts. This is particularly true for socio-economic aspects that have proven been relatively disregarded within the different frameworks. One of the main reasons mentioned was the technological angle that most of those frameworks were based on. Indeed, in most of the cases, they were meant to be implemented by and for the AEC industry for planning and design purposes. In this context, assessments are most likely to focus on aspects that they can grasp and tackle during this stage. Efforts must be done for the development of a more consensual framework. Moreover, there is room for a new type of framework that would consider in a greater extent the operational stage of an urban area. Such vision is re-enforced by the gain in performance and life span that well managed and maintained neighbourhoods demonstrate. However, this new vision brings some new requirements related to the framework dynamism. Undeniably, decisions taken at the operative stage are not much design-related but are closer to day-to-day managerial decisions, introducing a new time scale to the assessment. Real or near real-time assessment mechanisms are much needed in such context.

Additionally, there is a problem with local and temporal adaptability. In most cases, frameworks adapt themselves to different places and time based on a bespoke manner. They are no mechanisms that could allow the changes within the framework itself to automatically fit a particular situation. This issue relates to the weighing system used and how it can be changed. A new framework must open solutions in order to tackle the issue of adaptability.

Finally, transparency is another mentioned issue. There were a relatively few amounts of schemes that would describe in complete transparency the KPIs, their weights and the procedures to calculate them. It is essential to develop more inclusive and accessible schemes that allow all the parties in a good understanding of the assessment itself.

In other words, there is an important gap to be filled in the assessment of neighbourhood sustainability in real or near real-time. There is a need for a scheme that goes beyond the simple scoring and that serves as a decision support system for good operation and maintenance; a scheme that brings people into the loop and is meant to be more inclusive and cautious of the citizen satisfaction and city managers requirements.

Such vision is particularly relevant in the upcoming smart city paradigm where IoT and big data have the potential to records even the most complex aspects of the city ecosystem. Indeed, an efficient sustainability assessment and decision support tool is believed to leverage on the smart city paradigm, being at the crossroad of major technological assets such as IoTs, data mining, artificial intelligences and information systems. Some of the key technological assets considered for the development of the new framework have already been introduced in Chapter 2. Intelligent information systems such as semantic web tools, BIM or GIS coupled with cutting edge sensing technologies and sophisticated data processing such as forecast or data mining will empower the urban sustainability analysis with a better representation of the neighbourhood. It will improve the dynamism of such scheme by enabling a real or near-real-time assessment, recording previous states and even predicting future condition. A semantic approach is considered because of its ability to deal with heterogeneous data sources and systems. Information coming from the IoT to support the urban sustainability assessment will most likely come from heterogeneous sources, all with their own specification and structure. Semantic models are believed to unified sources around a common taxonomy. Furthermore, semantic web technologies enable the creation of linked data. They have the potential to link the assessment framework requirements to local policies. Therefore, no matter the place and time, the framework will have the ability to adapt its benchmarks automatically via the upload of a new policy. Finally, the development of a user-friendly web service with dashboards and a 3D interface will allow a more inclusive and transparent framework, easily accessible and understandable by everyone.

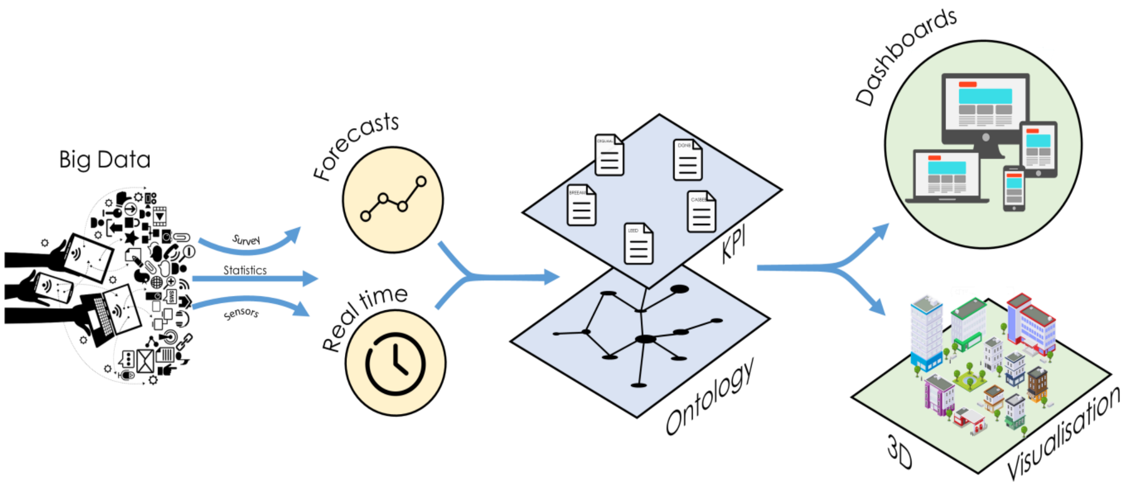


Figure 4-1 Urban sustainability assessment platform

Figure 4-1 presents the main features of the future platform. Data are collected from the IoT and then aligned with the urban sustainability assessment ontology to enhance their meaning, linking them to the built environment, agents and processes; and to define interconnections across scale and domains. Data processes can be introduced beforehand including clusterization, aggregation and disaggregation, scenario prediction and optimisation methods. A 3D interface based on BIM and cityGML technologies will enable user-friendly navigation and provides the semantic labelling of various components of the urban environment. Additionally, some dashboard will be integrated displaying the main outcomes of the assessment such as KPIs, real-time information, predictions, alerts and recommendations.

4.2 THE KPIS SELECTION.

4.2.1 Themes and Subthemes determination

Figure 4-2 Urban sustainability word cloud

The first step of the development of an urban sustainability assessment framework starts with the selection of indicators that best grasp environmental, social and economic requirements at the urban scale. It is a difficult task as no consensus has yet been found in the definition of sustainability. The literature review has pointed out that disperse key performance indicators were addressed within the current frameworks studied, often let at the local experts' regard. Since no model prevails, the selection of a specific framework is inadequate, especially in the future prospect of the creation of a semantic model. Consequently, a new framework is considered based on the already existing instances found in the literature. All the themes, sub-themes, criteria and indicators addressed within the 29 referred frameworks have been listed and compared [16], [17], [59]–[68], [50], [69]–[77], [51]–[55], [57], [58]. The word frequency has been investigated within the themes in order to draw a first set of themes as shown in Figure 4-2. The terms of "water, energy, management, environment, transportation, waste, innovation, site, materials, planning, design, health, economic, use and community" are the most cited. However, even if these terms draw a first sketch for the definition of the themes, it cannot be considerate as 100% reliable. Indeed, on the one hand, there are some terms that are analogous and that cover the same idea such as location and site or planning and design. On another hand, some themes are too wide (environment, design or social) or too narrow (education or tourism) to be considerate.

A second approach has been initiated in order to gather equivalent concepts and to clustered families of entities. Figure 4-3 is a mind map that refers to actual instances found in the current urban sustainability assessment schemes. This map gathers similar concepts together and shows the links between them as well as the level of abstraction. Note that this mind map has no vocation to be rigorous but rather aims at leading a reflexion in the categorisation of urban sustainability concepts. Consequently, the author has identified via this mind map 8 main themes and 25 subthemes eligible for the development of a new framework. Table 4-1 gives the different themes and subthemes identified in the literature.

Table 4-1 USA Themes and SubThemes

Themes	Sub-themes
Resources and Climate	Energy, water, waste, materials, outdoor environment
Land use and ecology	Landscape, Land use, Heritage
Urban Design	Transportation/mobility, access, public spaces, amenities and infrastructures
Health and well-being	Health, safety, community identity, quality of life
Governance	Engagement and maintenance
Innovation	Innovative technologies, innovative practices
Equity/diversity	Housing, fairness, cultural and social diversity
Resilient economy	Business, subsistence, costs

- *Resource and climate* refers to the various resources used for the operation of an urban area and their impact on the local environmental conditions. Energy, water, waste and materials quality and provision are considered. In term of environmental impact, pollution, emissions, noise or urban heat island is addressed.
- *Land use and ecology* includes the landscape typology as well as the native ecosystem preservation as a factor for sustainability. The respect of the natural heritage during land planning is a key element here.
- *Urban Design* relates to the functional planning of urban areas as a vector for social inclusion, improved access and mobility. It addresses the provision of efficient services and facilities within public spaces for a community centred vision.
- *Health and Well-being* aims to measure the quality of the health and the provision of care within the community as well as the feeling of safety and well-being. This includes the prevention of hazards, the satisfaction and efficiency of public services, the construction of a unique community identity and people relationship to their immediate environment.

- *Governance* refers to the extent of the public involvement into local decisions making as well as governmental institutions initiatives for the development of more sustainable and fair places e.g. provision of innovative services for the community.
- *Innovation* relates to the penetration of innovative technologies such as IoT and practices such as digitalised services within the urban area.
- *Equity and Diversity* covers social inclusion and diversity within a community, advocating for a greater support to “minorities”. Right to education, equity, fairness of institutions and cultural consideration are at the core of such vision.
- *Resilient Economy* relates to the job market and employability of the community as well as the affordability of fundamental products and services. Additionally, it looks at the economic health of local businesses and the costs of infrastructures and public services.

Inside these themes, 25 subthemes have been defined. Note that some of the subthemes present are sometimes addressed as criteria within the literature. The choice to consider them as subthemes has been motivated by the degree of abstraction and the fact that there is room for the definition of subclasses of those or criteria.

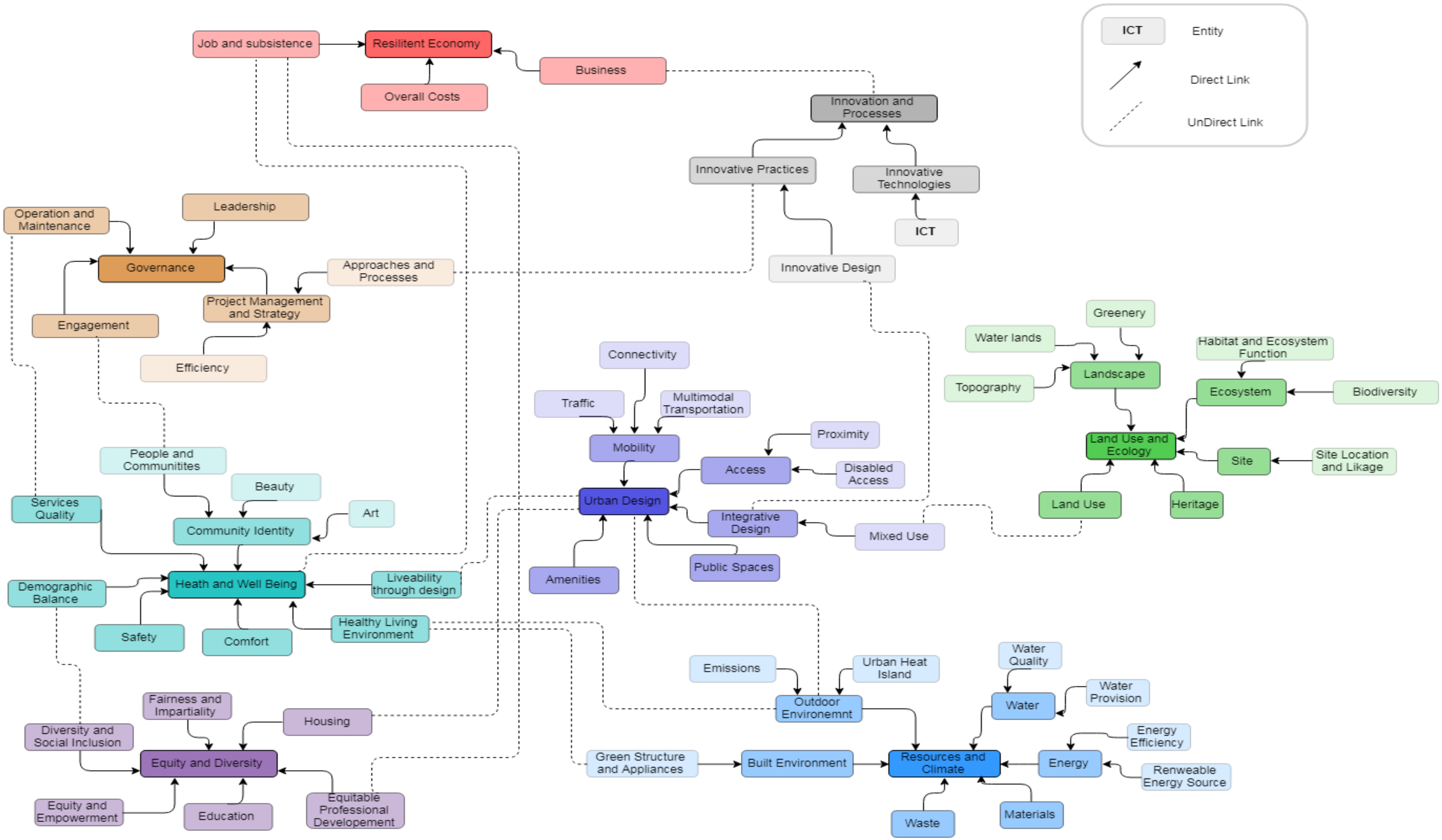


Figure 4-3 Themes and Subthemes mind map

4.2.2 Criteria Determination

The criteria have been identified after a second review of the frameworks. These criteria do not represent the entire diversity of criteria found in the literature but are those with enough occurrences to reasonably consider them as essential. Indeed, some frameworks with a particular focus such as PEER have too narrow and specific criteria that do not suit a more global approach. The criteria chosen were then restricted to a general and medium level of details to suit the framework purposes. Table 4-2 shows the criteria and the theme within they belong. For each subtheme a set of criteria has been selected in the literature. There are believed to be key aspects of the subthemes they refer to. A total of 90 criteria have been identified.

Those criteria are believed to form a solid basis for the definition of sustainability at the urban level. They cover various aspects from energy and water quality to equity in education and businesses economic health, covering land use, from brownfield to agricultural land, vegetation and natural species protection, public transport penetration or health care accessibility.

Table 4-2 Theme, sub-theme and criteria selection

Themes	Sub-themes	Criteria
Resources and Climate	Energy	Energy use
		Energy sources
		Infrastructure performance
		Energy losses
		District heating and cooling
	Water	Water use
		Water sources
		Stormwater
		Wastewater/sewerage
		Water quality / Sanitation
		Water leaks and losses
		Irrigation
	Waste	Waste treatment
		Hazardous waste
		Waste intensity
		Waste Collection
	Materials	Responsible sourcing
		Low impact materials
	Outdoor Environment	GHG
		Outdoor air quality
		Heat island effect
		Noise pollution
		Light pollution
Land use and ecology	Landscape	Streams, watersheds and floodplains
		Natural topography
	Land use	Nature-friendly design
		Agriculture
		Limit of growth
		Brownfields
		Construction development
		Consistency with upper-level planning
	Heritage	Vegetation
		Habitat
		Sensitive land
		Soil protection
Urban Design	Transportation/ mobility	Public transports
		Vehicles impacts

		Bicycle routes and amenities
		Pedestrian routes
	Access	Connected streets
		Accessible services/proximity
		Handicapped Accessibility
	Public Spaces	Recreation areas
		Adequacy
		Street multiple function / mixed-use
		Natural areas
	Amenities / Infrastructures	Parking
		Green design
		Infrastructure capacity
		Community services and facilities (hardware)
Health and well-being	Health	Healthy, local, affordable food
		Healthcare
		Pollutants
		Physical activities and healthy lifestyle
	Safety	Safe street/crime
		Natural risks
		Traffic safety
	Community identity	Beauty / inspiration
		Social network
		History preservation
		Public art
	Quality of life	Indoor environment
		Biophilic environment
		Hunger
		Demography
Governance	Engagement	Public involvements
		Raising awareness
		Local governance
		Management of facilities
Innovation	Innovative technologies	ICT
		Digital services
	Innovative practices	Performant practices / smart logistic
Equity/diversity	Housing	Housing type/balance
		Housing availability
	Fairness	Forced displacement
		Education

		Gender Equality
		Community services
		Internet connectivity
		labour standards
	Cultural and social diversity	Socio-economic diversity
		Cultural diversity
		Art and cultural facilities
Resilient economy	Business	Investments
		Economic growth
	Subsistence	job creation, employment
		Affordability
		Poverty
		Decent work
	Costs	Lifecycle Costs

4.2.3 Indicators Determination

The review of the frameworks allowed us to determine the most commonly used indicators as well as criteria. The selection of the indicators has been done while keeping in focus the main purpose of this new framework and a realistic implementation. Thus, some indicators such as "number of properties with pesticide-free or Integrated Pest Management plans to reduce toxicity " [56] have not been taken into account because too specific or hard to be metered. Additionally, some indicators seemed redundant and therefore have been excluded from the KPIs. From the original 304 indicators considered, 193 distinct ones have been selected from the different frameworks.

Table 4-3 presents the list of those indicators along with a short description and the unit of measure associated.

Those indicators have been the object of several static assessment schemes. There is a need to investigate means to measure those in real or near real-time for a realistic implementation at the operative stage of an urban area. This aspect is further developed in the following Section 0.

Despite a rigorous approach for the KPIs selection based on their occurrences within the literature, this method does not prevent a certain subjectivity in their selection. Indeed, in the 304 indicators originally selected, some were redundant in nature and the author had to choose the most suited ones. In order to end up with a valid framework, those indicators will be subjects of a consultation across experts to objectively determine which ones are the most relevant for a near-real-time implementation. A DELPHI consultation involving various domain experts will be carried out. The survey will question the overall relevance of certain indicator for sustainability assessment and will go further by questioning their potential relevance within a near real-time framework. Indeed, the vision of a dynamic framework that can be applied at the operation stage of development brings new considerations (especially the temporal dimension of certain KPIs) that need to be discussed. The USA framework validation is presented in Section 4.4.

Table 4-3 Indicators selection

Indicator	Description	Unit
Energy		
Total Energy Demand	Energy demand of buildings	kWh/Time unit
Solar Potential On Roof	Potential energy generated if PV were installed	kWh/m²
Electricity From Renewable Sources	Share of electricity from renewable sources	%
Heating From Renewable Sources	Share of heating from renewable sources	%
On-Site Generation	Share of on-site generation	%
Reliable Electricity	Share of the population using reliable electricity	% of population with
Electrical Losses	Electrical loss in the electrical network	kWh/Time unit
Heat Losses	Heat losses in the distribution network	kWh/Time unit
Water		
Potable Water Demand	Demand and consumption of potable water	l/Time unit
Crop Water	Crop water productivity (tons harvested per unit irrigation water)	tonneneharvested/l
Potable Water Supply	Share of the population supplied with potable water	% of population with
Potable Water Resources For Landscape Irrigation	Use of potable water resources for landscape irrigation	l/Time unit
Water Recycled	Share of water recycled	%
Runoff Treated Or Retained	Share of runoff from impervious areas within the site that can be treated or retained	%
Ratio Of Impervious To Pervious Area	Ratio of impervious to pervious area that accommodates stormwater infiltration	%
Waste Water Treatment	Share of wastewater being treated with at least primary treatment	%
Waste Water Collection	Share of the population served by wastewater collection	% of population with
Temperature Of Creek Runoff	Temperature of creek runoff	degree C
Drinking Water Minimum Quality	Share of drinking water meeting minimum quality	%
Water Runoff Quality	Quality of stormwater run-off (pollutant concentration)	pollutant concentration

Water Leaks	Number of water leaks within the network over a period	Units/Time unit
Waste		
Waste Recycled	Share of waste recycled (included compostable)	%
Local Composting unit	Provision of local composting	unit/capita
Organic Waste For Energy Generation	Share of organic waste for energy generation	%
Hazardous Waste Generation	Hazardous waste generation per capita	tonne/capita/Time unit
Hazardous Waste Recycled	Share of hazardous waste recycled	%
Waste Generation	Share of solid waste	tonne/capita/Time unit
Regular Waste Collection	Share of the population with regular waste collection	% of population with
Garbage Separation	Garbage separation	yes/no
Waste Collection Point	Share of the population within 200m from a collection point	% of population with
Frequency And Capacity Of Collection	Frequency and capacity of collection	tonnes/Time unit
Materials		
Materials From Local Sources	Share of material from local sources	%
Recycled Materials	Share of recycled materials	%
Cut And Fill Materials	Share of materials from cut and fill	%
Outdoor Environment		
GHG Emissions	GHG emissions	CO_2e/Time unit
Concentration Of NOx And Particles	Concentration of NOx and particles	ppmv
Air Quality Level	Share of time with air quality minimum level	% of Time
Thermal Gradient Differences	Thermal gradient differences between urban and rural area	degree C
Open Spaces	Share of open spaces	% (surface ratio)
Shaded Area Of Public Space	Share of shaded area of public space	% (surface ratio)
Ambient Noise	Ambient noise	dB(A)
Solar Reflectance Index	Solar Reflectance Index	SRI average on roof
Outdoor Light Glare	Outdoor light glare	UGR (unified glare rating)
Light Trespass	Light trespass	lux

Skyglow	Light in the night sky (skyglow)	NELM (Naked Eye Limiting Magnitude)
Landscape		
Coastal And Marine Area Protected On Site	Share of coastal and marine area protected	% (surface ratio)
Wetlands Streams And Shoreline Buffers Loss	Share of wetlands, streams, and shoreline buffers net-loss	% (surface ratio)
Transformed Land Topography On Site	Share of land topography transformed	% (surface ratio)
Area Slope	Area slope	degree
Land use		
Tree Planted	Number tree planted per capita over a time period	unit/ Time unit
Corridor For Biodiversity Connectivity	Presence of corridor for biodiversity connectivity	yes/no
Working Lands Good Management Practices	Share of working lands with good management practices	% (surface ratio)
Prime Agricultural Preserved On Site	Share of the site located on prime agricultural preserve for food production	% (surface ratio)
Infill Development	Share of infill development	% (surface ratio)
Brownfield Greyfield Redevelopment	Share of new development on brownfield and/or greyfield	% (surface ratio)
Existing Infrastructure Redevelopment	Share of new relying on existing infrastructure	%
Built-Up Area By Sector On Site	Share of the total built-up area by sector	% (surface ratio)
High-Risk Areas	Share of population living in designated high-risk areas	% of population
Heritage		
CO2 Sequestered By Vegetation	Measures CO2 sequestered by vegetation	CO2e/Time unit
Native And Non-Native Plants	Amount and diversity of native and non-native plants in area	unit
Acreage Of Area Covered By Vegetation On Site	Share of area covered by vegetation	% (surface ratio)
Diversity Of Species	Amount and diversity of species in area	unit
Native Species	Share of change in number of native species	%
Land Capability Classification	Land Capability Classification	
Reinforcement Or Re-Grading Of Slopes	Presence of soil protection by reinforcement or re-grading of slopes	yes/no
Transportation		
Transit Frequency	Average time between two oncoming public transport	minute

Public Transport Penetration	Public transport penetration rate	km/km²
Number Of Passengers	Number of passenger over a time period	unit/Time unit
Vehicle Miles Travelled	Vehicle distance travelled over a time period	km/Time unit
Diesel Emissions	Diesel emissions over a time period	CO2e/Time unit
Carbon Emissions	Carbon emissions over a time period	CO2e/Time unit
Personal Automobiles	Number of personal automobiles per capita	unit/capita
Energy Efficient And New Fuel Vehicles	Share of energy-efficient and new-fuel vehicles	%
Eco-friendly Refuelling Stations	Number of Eco-friendly Refuelling Stations	unit/capita
Bicycle Paths And Lanes Per Capita	Bicycle paths and lanes penetration rate	km/km²
Bicycle Racks And/Or Storage	Number of bicycle racks and/or storage per housing units	unit/capita
Bicycle Paths And Lanes Wide	Average wide of bicycle paths and lanes	m
Entire Pedestrian Area	Share of pedestrian areas	% (surface ratio)
Side Walk Penetration	Share of roadway with sidewalk	% (length ratio)
Side Walk Wide	Average wide of sidewalk	m
Modes Of Transportation	Share of the population using different modes of transportation	% of population with
Off-road Trail Proximity	Share of households within 3 miles off-road trail	% of population with
Intersections Nearby	Existing number of intersections per square mile nearby	unit/km²
Proximity and Access		
Services And Facilities Proximity	Share of the population living within 0.5 miles of services and facilities	% of population with
Accessible Crosswalks	Share of accessible crosswalks	%
Accessible Transit Facilities	Share of accessible transit facilities	%
Public Space		
Recreational Area	Share of recreational areas (indoor and outdoor)	% (surface ratio)
Recreational Area Satisfaction	Share of population satisfied with recreational areas	% of population with
Homogeneity Of Housing Design	Presence of homogeneous housing design	yes/no
Building Height To Street Width Ratio	Building-height-to-street-width ratio	%
Garage Doors Street Length Ratio	Garage doors along street ratio	%

Mixed Use Space	Share of total area land zoned for mixed use	% (surface ratio)
Natural Areas Proximity	Share of the population living within 0.5 miles of natural areas	% of population with
Natural Areas Penetration	Share of natural areas	% (surface ratio)
Natural Area Satisfaction	Share of population satisfied with natural areas	% of population with
infrastructures		
Parking Slots	Number of allocated parking slots	unit/capita
Carpooling Slots	Number of allocated carpooling slot	unit/capita
Parking Proximity To Public Transport	Parking lot proximity to public transport	m
Open Air Parking With Multiple Use	Share of open-air parking with flexible use, e.g. market stalls, play areas	%
Rooftop Greening	Share of rooftop greening	% (surface ratio)
Wall Greening	Share of wall greening	% (surface ratio)
Buildings With Relevant Sustainability Certification	Share of buildings with relevant sustainability certification	%
Efficient Building Envelope	Average efficiency of building envelope	U-value
Internal Buildings Area	Internal buildings area	m^2
Public Appliances With Good Performances	Share of public appliances with good environmental performances	%
Health		
Healthy Food Store Proximity	Share of the population living within 1/4 miles to store proposing healthy food	% of population with
Fresh Food Produced Through Local Agriculture	Amount of fresh food produced through local agriculture	tonnes
Urban Food Desert	Share of the population living in an urban or rural food desert	% of population with
Healthy Food Public School	Presence of fresh fruits and vegetables in the largest public school	yes/no
Growing Spaces For Fruits And Vegetables	Growing Spaces for fruit and vegetable (sq. m. per dwelling unit)	m^2/dwelling
Effective Financial Protection For Health Care	Share of the population with effective financial protection for health care	% of population with
Pollutant Loads	Pollutant loads over a time period	kg/Time unit
Healthy Weight	Share of the population with healthy weight	% of population with
Leisure Time Physical Activity	Share of adults aged 20+ with leisure-time physical activity	% of population with
Public Schools Physical Activity	Share of public schools that require some form of physical activity	%
Safety		

Safety Feeling	Share of the population feeling safe	% of population with
Fires	Number of fire per capita over a time period	unit/capita/Time unit
Crimes Against Property	Number of crimes against property over a time period	unit/capita/Time unit
Violent Crimes	Number of violent crimes over a time period	unit/capita/Time unit
Infrastructure On Flood Area	Share of infrastructure on flood area	%
Flooding Risk Area On Site	Share of flooding risk area	% (surface ratio)
Police Officers	Number of police officers per capita	unit/capita
Response Time Of Police Department	Response time of police department after call	minute
Health Care Practitioners	Health care practitioners per capita	unit/capita
Firefighter	Number of firefighter per capita	unit/capita
Reponses Time Of Emergency	Reponses time of emergency after call	minute
Road Traffic Deaths	Road traffic deaths per capita	unit/capita/Time unit
Roadways Limited To Low Speed	Share of roadways limited to low speed	% (length ratio)
Community identity		
Residents Volunteering	Share of residents who volunteered over the past 3 years	%
Local Community Groups/Events Investment	Sponsor, facilitate and/or provide local community groups/events	currency
Historical Building Recovered	Share of recovery or reconversion	%
Urban Art Pieces	Number of urban art pieces	unit/km²
Quality of life		
Complaint Regarding Air Quality	Number of complaint regarding air quality over a time period	unit/capita/Time unit
Natural Ventilation	Share of natural ventilation	% (Volume Ratio)
Efficient HVAC	Average HVAC efficiency	EER (Energy Efficiency Rating)
Complaint Regarding Thermal Comfort	Number of complaint regarding thermal comfort over a time period	unit/capita/Time unit
Timeout Thermal Comfort Levels	Share of time out thermal comfort levels	% of Time
Low Volatile Organic Compound (VOC)	Low Volatile Organic Compound (VOC) concentration	ppmv
Satisfaction Of Services	Share of population satisfied with services	% of population with
Urban Farm	Urban farm superficies	m²

Minimum Level Of Dietary	Population under the minimum level of dietary	unit of people below
Population Density	Population density	unit/m²
Demographic Projection	Demographic projection	unit/m²
Local governance		
Local Officials Elected To Office	Number of local officials elected to office	unit
Community Grants Programs	Number of community grants programs	unit/capita
Innovation		
Monitoring Systems	Number of monitoring systems	unit/building
Equipped Buildings	Share of equipped buildings	%
Communications Infrastructure Speed	Communications speed	Mbps
Housing		
Social Housing	Share of social housing	%
Housing Penetration	Housing penetration	unit/km²
Fairness		
Relocated Resident Or Business	Number of relocated resident or business due to a project	unit/capita/Time unit
Student Completing Primary Education	Share of student completing primary education	%
Student Completing Secondary Education	Share of student completing secondary education	%
Primary Student-Teacher Ratio	Primary student/ teacher ratio	%
High Education Degree	Number of high education degree per capita	unit/capita
Enrolment Rates	Enrolment rates	%
Public School Students With Reading Proficiency	Share of third-grade public school students meet or exceed reading proficiency	%
Enrolment Rates For Women And Men	Enrolment rates for women and men	%
Employment Rates For Women And Men	Employment rates for women and men	%
Women Elected To City-Level Office	Share of women elected to city-level office	%
People With Broadband Access	Share of population with broadband access	% of population with
Civil And Human Rights Complaints	Civil and human rights complaints over a time period	unit/capita/Time unit

Social and cultural diversity		
Diversity Index	Diversity Index	
Income Distribution	Income distribution curve (Lorentz curve and Gini coefficient)	%
Events Celebrating Social And Cultural Diversity	Number of events celebrating social and cultural diversity	unit/Time unit
Boards And Commissions Reflect The Ethnic Diversity	Share of local board and commission reflecting the ethnic diversity of the community	%
Cultural Institution Proximity	Share of the population living within 1/5 mile from cultural institution	% of population with
Business		
Governmental Investments	Share of governmental investments	%
Funds Deposited In Locally Owned And Operated Financial Institutions	Increase of the total funds deposited in locally owned institutions over time	currency/capita
Commercial Lease/Vacancy Rates	Commercial Lease/Vacancy Ratio	%
Businesses Penetration	Number of businesses per capita	unit/capita
Valued Companies	Share of valued companies (gross sales > x)	%
Business Establishments	Number of business establishments over time	unit/Time unit
Sales From Businesses	Sales from businesses over time	currency/capita
Economic Projection	Economic projection	currency/time unit
Subsistence		
Employment By Sectors	Total employment in targeted industry sectors over time	unit/Time unit
Unemployment Rates	Unemployment rates	%
Job Housing Ratio	Job/housing ratio	unit/unit
Transport Affordability	Transport affordability	currency
Housing Affordability	Share of total housing designated as “affordable”	currency
Food Affordability	Food affordability	currency
Art And Cultural Event Affordability	Art and cultural event affordability	currency
Healthcare Affordability	Healthcare affordability	currency
Electricity Affordability	Electricity affordability	currency
Gas Affordability	Gas affordability	currency

Water Affordability	Water affordability	currency
Telecommunication Internet Affordability	Telecommunication/internet affordability	currency
Self-sufficient People	Share of self-sufficient population	%
Homeless Population	Share of homeless population	%
Population Living In Informal Settlements	Share of population living in informal settlements	%
Hours Spent On Paid And Unpaid Work	Average number hours spent on paid and unpaid work	h
Low Incomes Population	Share of people with incomes below 50% of the median income	%
Costs		
Development Cost	cost-effective development	currency

Note that at this stage of development, the consideration of certain indicators does not guaranty they can be measured in real time. When considering real time assessment, one must question the meaning of real time for each indicator. Indeed, the end goal of monitoring certain aspects is to track changes but those changes will happen at different paces. Consequently, defining "real time" amounts to understand at which pace those changes occur for a particular phenomenon. The following sections will further develop this question via the investigation of KPIs determination methods and the DELPHI consultation.

4.3 THE KPIS DETERMINATION METHOD

A realistic implementation of the indicators requires the investigation of mean of measurement. Indeed, in the perspective of a real or near real-time operation of the assessment, technological assets that can be effective in real-time are needed. In this section, a systematic review of technologies that could support the measurement and calculation of the KPIs in real-time is done. Table 4-4 gives an overview of the type of technologies that can be used for the KPIs determination. A set of references are also given in some cases giving real instances to support the choice of a pair technology-KPI. Those technologies are aligned with the ones presented in Section 2.2. Those are the core technologies for the development of the smart city such as sensors and meter from the IoT, BIM and GIS information system, open government platform, and crowdsourcing and web data mining.

- *Sensors* are often related to graspable physical phenomena such as temperature, pressures and loads, humidity, energy transfers, volumes, weights and flows, noises, lights power or elements concentration. They can also measure events and movements including, for example, video cameras, smart fire alarms, smart security alarms, GPS and RFID. Therefore, every indicator relating to those aspects such as energy, water, materials, pollutant concentration, transportation, alarms and alerts about fire and crime or divers events can be measured via various types of sensors.
- *Remote Sensing* is technically a form of sensing method and therefore relates to sensors. However, its unique use when combined with *GIS technologies* has motivated the creation of a distinct category. Remote sensing includes assets such as satellite images, high-resolution aerial pictures, and LiDAR. The content they produce along with intelligent "object" detection algorithms are the basis for the creation of accurate GIS that can serve as a valuable source of information for the

determination of KPIs. In the present case, the information does not necessarily come from the sensor directly but from a frequently up to date GIS. Remote sensing and GIS technologies most often intervene for the definition of geospatial information from identification of areas to localisation of objects, distances between objects, slopes, water and vegetation coverage etc. It is a key technology for the determination of the functional layout of a city, catching aspect such as mixed-use, natural area and transport network penetration, urban sprawl, proximity to services and facilities, etc.

- *Statistics and Open Government Data* relates to information that can solely be gathered by governmental authorities. Statistical insights about public services such as the enrolment rate in public education, hospitalisation motives, transport, water and energy prices as well as garbage collection information (if provided by the public sector) can be monitored and made available via open governmental platform API. Additionally, the public sector is the only actor able to provide demographic information (with census campaign for instance) and socio-economic insights with income and businesses taxes.
- *BIM technologies* give information about the various building elements and appliances present within an urban area. Building meta-information about geometries, materials, efficiency and costs can be retrieved as well as information on occupancy, shading effects and machines provision. It can help in the determination of KPIs relating to the quality of the built environment.
- Finally, *Crowdsourcing and Data Mining* leverages on citizens and businesses data provision for the detection of events, defects and to sense satisfaction and well-being. Some example includes social media data mining to evaluate the degree of satisfaction over certain public services, voluntary information delivery from people and businesses information via open platforms and objects detection via mobile apps.

Table 4-4 USA means for measurements

Technology	Indicator			
Sensor	Total Energy Demand		Transit frequency	[539]
	Electricity From Renewable Sources		Number Of Passenger	[540], [541]
	Heating From Renewable Sources		Vehicle Miles Travelled	[542], [543]
	On-Site Generation		Modes Of Transportation	
	Electrical Losses	[544]	Personal Automobiles	[542], [543]
	Heat Losses	[545]	Vehicle Diesel Emissions	[542]
	Reliable Electricity		Vehicle Carbon Emissions	[542]
	Potable Water Demand,		Parking Slots	[136]
	Potable Water Resources For Landscape Irrigation,		Carpooling Slots	[136]
	Water Recycled,		Public Appliances With Good Performances	
	Runoff Treated Or Retained,		Pollutant Loads	
	Waste Water Treatment,		Police Officers	[546]
	Drinking Water Minimum Quality	[547]	Firefighter	[546]
	Water Leaks	[148]	Response Time Of Police Department	
	Temperature Of Creek Runoff		Response Time Of Emergency	
	Water Runoff Quality	[547]	Fires	[548]
	Runoff Treated Or Retained		Crimes Against Property	
	Waste Recycled	[152]	Violent Crimes	
	Organic Waste For Energy Generation	[152]	Complaint Regarding Thermal Comfort	[549], [550]
	Hazardous Waste Generation	[152]	Low Volatile Organic Compound (VOC)	
	Hazardous Waste Recycled	[152]	Fresh Food Produced Through Local Agriculture	[150]
	Waste Generation	[152]	Timeout Thermal Comfort Levels	
	Local Composting unit	[551]	Natural Ventilation	
	Thermal Gradient Differences		Efficient HVAC	
	Ambient Noise		Communications Infrastructure Speed	
	Solar Reflectance Index	[552]	Monitoring Systems	
	Outdoor Light Glare	[553]	Equipped Buildings	
	Light Trespass		Civil And Human Rights Complaints	[554]
	Skyglow	[555]	People With Broadband Access	
	Concentration Of NOx And Particles	[556]	Area Slope	

	Air Quality Level		Native And Non-Native Plants	[557]
	Population Density	[558]	Diversity Of Species	[557]
			Native Species	[557]
			Native Species	[557]
	Solar Potential On Roof	[559]	Bicycle Paths And Lanes Per Capita	
	Runoff Treated Or Retained	[560]	Bicycle Paths And Lanes Wide	[561]
	Crop Water	[351], [562], [563]	Entire Pedestrian Area	[561]
	Ratio Of Impervious To Pervious Area	[560]	Side Walk Penetration	[561], [564]
	Local Composting unit	[565], [566]	Side Walk Wide	[561], [564]
	Waste Collection Point	[565], [566]	Intersections Nearby	[561], [564]
	GHG Emissions	[567]	Off-road Trail Proximity	
	Concentration Of NOx And Particles	[568]	Eco-friendly Refuelling Stations	[564]–[566]
	Open Spaces	[561]	Bicycle Racks And/Or Storage	[564]–[566]
	Shaded Area Of Public Space	[268]	Services And Facilities Proximity	
	Wetlands Streams And Shoreline Buffers Loss	[569]	Accessible Crosswalks	[564]–[566]
	Transformed Land Topography On Site		Accessible Transit Facilities	[564]–[566]
	Area Slope		Recreational Area	[561]
	Prime Agricultural Preserved On Site	[349]	Building Height To Street Width Ratio	
	Infill Development	[570]	Garage Doors Street Length Ratio	
Remote Sensing + GIS	Brownfield Greyfield Redevelopment		Mixed Use Space	[561]
	Built-Up Area By Sector On Site		Natural Areas Penetration	[571]
	High-Risk Areas	[572]–[574]	Natural Areas Proximity	
	Corridor For Biodiversity Connectivity		Parking Slots	[564]–[566]
	Acreage Of Area Covered By Vegetation On Site	[571]	Carpooling Slots	[564]–[566]
	Tree Planted	[571]	Open Air Parking With Multiple Use	
	Native And Non-Native Plants	[557]	Parking Proximity To Public Transport	
	Diversity Of Species	[557]	Healthy Food Store Proximity	
	Native Species	[557]	Urban Food Desert	
	CO_2 Sequestered By Vegetation	[352]	Health Care Practitioners	
	Public Transport Penetration		Growing Spaces For Fruits And Vegetables	
	Infrastructure On Flood Area	[572]	Urban Art Pieces	[565], [566]
	Flooding Risk Area On Site	[572]	Urban Farm	
	Roadways Limited To Low Speed		Social Housing	
	Cultural Institution Proximity		Housing Penetration	
	Business Establishments		Population Living In Informal Settlements	[575]
	Businesses Penetration			

Statistics and Open Government Data	Reliable Electricity	[576]	Relocated Resident Or Business	
	Waste Water Collection		Enrolment Rates	[401], [412]
	Frequency And Capacity Of Collection		Enrolment Rates For Women And Men	[401], [412]
	Regular Waste Collection		Employment Rates For Women And Men	[401], [412]
	Garbage separation		Women Elected To City-Level Office	[401], [412]
	Recycled Materials	[577]	Boards And Commissions Reflect The Ethnic Diversity	[401], [412]
	GHG Emissions		Income Distribution	[401], [412]
	Working Lands Good Management Practices		Governmental Investments	[578]
	Reinforcement Or Re-Grading Of Slopes		Valued Companies	[401], [412]
	Vehicle Diesel Emissions		Sales From Businesses	[401], [412]
	Vehicle Carbon Emissions		Employment By Sectors	[401], [412]
	Road Traffic Death		Unemployment Rates	[401], [412]
	Violent Crimes		Job Housing Ratio	[401], [412]
	Healthy Food Public School		Transport Affordability	[401], [412]
	Effective Financial Protection For Health Care	[129], [138], [579]	Electricity Affordability	[401], [412]
	Healthy Weight	[129], [138], [579]	Heat Affordability	[401], [412]
	Local Community Groups/Events Investment	[578]	Water Affordability	[401], [412]
	Population Density		Telecommunication Internet Affordability	[401], [412]
	Minimum Level Of Dietary	[129], [138], [579]	Self-sufficient People	[401], [412]
	Local Officials Elected To Office		Hours Spent On Paid And Unpaid Work	[401], [412]
	Community Grants Programs	[578]	Low Incomes Population	[401], [412]
	Student Completing Primary Education	[401], [412]	Primary Student Teacher Ratio	[401], [412]
	Student Completing Secondary Education	[401], [412]		
	High Education Degree	[401], [412]		
	Public School Students With Reading Proficiency	[401], [412]		
BIM	Reliable Electricity		Buildings With Relevant Sustainability Certification	
	Waste Water Collection		Efficient Building Envelope	
	Local Composting unit		Rooftop Greening	
	Materials From Local Sources	[254], [580]	Wall Greening	
	Recycled Materials	[577], [580]	Internal Buildings Area	
	Cut And Fill Materials	[254], [580]	Historical Building Recovered	
	Shaded Area Of Public Space	[268]	Monitoring Systems	
	Homogeneity Of Housing Design		Equipped Buildings	
	Development and Operation Costs	[168]		

Crowdsourcing and Data Mining	Native And Non-Native Plants	[557]	Complaint Regarding Air Quality	[581]
	Diversity Of Species	[557]	Satisfaction Of Services	[554], [582]
	Native Species	[557]	Civil And Human Rights Complaints	[554]
	Modes Of Transportation		Events Celebrating Social And Cultural Diversity	
	Energy Efficient And New Fuel Vehicles		Housing Affordability	[583]
	Recreational Area Satisfaction	[582]	Electricity Affordability	
	Natural Area Satisfaction	[582]	Heat Affordability	
	Healthy Food Public School	[582]	Water Affordability	
	Leisure Time Physical Activity	[584]	Telecommunication Internet Affordability	
	Safety Feeling	[585]	Homeless Population	[586], [587]
	Residents Volunteering			

Overall, there are good evidences that all the indicators defined for the USA framework can be supported by one of those technologies. The smart city is likely to be an enabler for the development of real-time sustainability assessment and reliable decision support systems. However, some concerns remain toward the concept of real-time in the present framework. Indeed, the “real-time” concept is flexible and changes from an aspect to another. It relates to the extent of change of a certain phenomenon over a time period. For instance, energy flows are considered as an instantaneous phenomenon that requires frequent measurements while urban sprawl is a really slow process where a monthly capture could be reasonably considered as real-time. This introduced the question of which time scale is relevant for a certain KPIs and whether or not slow phenomena are even relevant in the context of a real-time assessment. Consequently, an expert consultation has been considered for the validation of the KPIs for the construction of a coherent framework.

4.3.1 Note on the weighting system

Section 2.1.3.3 of the literature review has investigated the link between weighting systems of the available urban sustainability assessment schemes and local adaptability. Indeed, studies have demonstrated that weighting changed in function of specific issues to stress in certain region [93]. For instance, water indicators will hold more importance in region with water scarcity as the issue is considered of great importance. Consequently, setting weights and credits must be done relatively to the place, time and culture in which the assessment is realised.

The author originally intended to address the weighting system feature within the DELPHI consultation however, in the light of such perspective it is irrelevant to do so as this one will only be valid for a certain place and time. Instead of focusing on the weight to be set, an interesting approach would be to considered means to capture which aspects are the most important automatically or semi-automatically. For instance, one could integrate in web-based assessment services a feature that track the aspects the most commonly queried by the users or such information could be retrieved from authoritative open platform API or linked data. Future work should focus in the implementation of solutions to tackle adaptability of the weighting system according to different places in an efficient manner.

4.4 FRAMEWORK VALIDATION: THE DELPHI CONSULTATION

4.4.1 Background and objectives

Summing up what stated in Section 3.4.5, the DELPHI survey has been designed in order to validate the urban sustainability framework developed. Experts from different domains such as energy systems, smart cities governance, transport, big data, sustainability have been contacted to evaluate the degree of relevance of the KPIs as well as the feasibility of including them into a real-time assessment scheme. The consultation has integrated several stages including two pilot versions that have served for the creation of the final questionnaire; and two consultation rounds on this same questionnaire.

4.4.2 Expert panel selection

Because the consultation requires deep knowledge across different domains, recruiting the right experts is essential for the relevance of the study. More, the expert selection must follow specific criteria in order to avoid biases. The researcher must agree on criteria such as gender, professional experience, education, employment or designation before selecting experts [588]. It is recommended to ask experts from various fields of expertise and location to minimize biases and context-specific issues [517], [588], [589].

Concerning the panel size, there is no evidence regarding the optimal panel size [590]. However, it is often recommended to gather between 20 and 50 experts across different fields of expertise [517], [590], [591].

The expert selection has been done on the basis of their knowledge in various areas. Research articles on urban sustainability, urban energy performance indicators, smart cities, ICTs, sustainability assessment schemes etc have been investigated and authors were contacted when the content appeared to be valuable for the survey. The expertise has been evaluated based on both quantitative and qualitative aspects of their articles publications in their respective domains. Overall, on 190 experts approached, 32 answered for a participation ratio of 16.8%. Internationally distributed, the experts are all part of research organisms in recognized universities or companies (Figure 4-4). McKenna and Jirwe et al. highly recommend to inform the experts on the specific focus and purposes of the research conducted [590], [592]. To this end, each expert was contacted by email which included a section on the scope of the study. The experts were informed of the other participants although experts' anonymity was preserved to avoid biases. Additionally, 10 "insider experts" have been asked to answer to the pilot versions of the survey in order to get feedback on the survey's structure itself. Those experts were participants of the different projects mentioned in section 3.4.4 and were contacted within the frame of the projects to collaborate in the survey development.

Figure 4-4 Distribution map of the participant

4.4.3 Pilot versions

The dynamic, near real-time and operative dimension of the new framework brings new issues that need to be discussed. Indeed, it is interesting to know whether the evaluation of certain indicators is feasible and appropriate in near real-time and the extent to which it would affect decision making, especially operational and managerial.

A pilot version of the survey has first been drawn and distributed to a small panel of experts in order to assess the quality of the survey itself. The first version was divided into the 8 main themes addressed in the framework namely:

Resources and climate (45 indicators): related to Energy, Water, Waste, Materials, Emissions and Outdoor environment

Land use and ecology (20 indicators): related to Landscape, Land Use and Nature Heritage

Urban Design (40 indicators): related to Transport, Access and Proximity, Public Space and Infrastructure

Health and Well-Being (38 indicators): related to Health, Safety, Community Identity, Quality of Life

Governance (2 indicators): Local Governance and Engagement

Innovation (3 indicators): Innovative technologies and practices

Equity and Diversity (19 indicators): Housing, Fairness, Cultural and Social Diversity

Resilient Economy (26 indicators): Business, Subsistence and Costs.

The survey was questioning the relevance, measurability and improvability of the KPIs in a YES/NO format where the respondent would choose if a specific KPI is irrelevant, hardly measurable and/or not improvable. The experts were also asked to rank their knowledge in each theme in order to ponder their answer in the previous questions.

The first version was opened from February 21st 2017 to August 31th 2017. Over that period, on 28 respondents that started answering the survey, only 4 finished it, for a response rate of 14.3%. Issues have been raised by many experts concerning the length of the survey. Indeed, the survey was designed to answer the 3 questions on each 193 KPIs which results in a considerable effort from the experts. This type of intensive questionnaire did not comply with the voluntary aspect of the survey. Therefore a second pilot version has been designed.

In this second version, KPIs have been clustered based on similar features and mean of measure. For instance, all energy consumption and production-related indicators have been grouped and are questioned on their relevance, measurability and improvability. From the 193 KPIs initially addressed, 71 clusters have been created and questioned. Additional, some new features were integrated to the survey designed that allowed the expert to "jump" the section in which they estimated their knowledge was not sufficient enough which ultimately would shorten the time required to answer the survey. The new pilot version was opened from July 20th 2017 up until December 31st 2017. For this survey, experts from the projects and the literature have been approached in order to gather a greater amount of expert and catch a more significant response rate. Therefore, out of 57 respondents, 6 answered the survey for a response rate of 10.5%. Despite the reduced amount of questions and the new survey design, the survey remained too long to be answered on a voluntary basis.

4.4.3.1 *Voluntary expert survey and incentives*

The pilot versions' low response rate raised an interesting issue that is the participation in a survey and the possible effect of incentives. In the past decade, the field of social sciences has seen a decline in the surveys' response rate [593]–[595]. This is especially true for online surveys where response rates are generally lower than face-to-face or postal survey [596], [597]. Therefore, strategies to significantly increase the response rates have been investigated [598]–[600]. The most common type of incentives that consistently and substantially improve the response rate ever since the mailed questionnaires is the monetary incentives [601]. Additionally, questioning a large set of individuals is a sure way to gain responses [599]. Overall, research has identified three principal reasons for participation: (1) altruism (fulfilling a social obligation), (2) survey-related (interest in the survey's topic) and (3) egoistic reasons (money, self-interest) [598]. If monetary incentives have proven efficiency to increase the response rate, another common feature is the will to take part in the results of a research [602]. This is especially true in the case of expert consultation where the idea to be integrated into a research overcome any monetary incentives [603]. Therefore, expert motivation to respond to a survey is more driven by the survey related than egoistic reasons. In the case of web-based expert survey, the principal factor for attrition seems to be related to the efforts needed to complete it [603]. A set of actions are recommended in order to increase the number of responses [603]: the survey layout should be designed to be time-efficient and subdivided into sections, closed-ended questions should be favoured over open-ended questions, the English language should be preferred for international reach, invitation letter should be personalised so that the respondents feel that they have been selected based on their achievements and knowledge, the survey provenance should be linked to a recognized institution such as a University, the purpose of the study should be clearly exposed, reminders should be sent cautiously as too many reminders tends to irritate and lower responses quality, respondents privacy should be preserved, a full transparency should be given concerning the data subsequent usage and the respondents should be given the opportunity to follow up the results of the research undertaken.

4.4.4 The Survey

Following the low answer rate of the pilot versions, a new approach has been considered. The main barrier for the expert appeared to be the length of the survey and the efforts needed to answer it which is incompatible with the voluntary aspect of the survey. Therefore, a subset of KPIs has been carefully selected among the 193 original indicators for validation. Those indicators have been chosen based on the technical ability to measure them. Indeed, the author has foreseen possible cases studies from the different research projects where KPIs could be calculated. The pilot site “The Work” in Ebbw Vale presented in Section 3.4.7.1 has been studied. Table 4-5 gives the data available in the pilot site of “The Work” and the 15 KPIs to which they relate. Additionally, the *EquippedBuildingsIndicator* and *MonitoringSystemsIndicator* that are based on the number of sensors installed can be determined for a total number of 17 KPIs addressed.

Table 4-5 The Work Data vs KPIs

Available Data	Provenance	Air Quality Level / Complaints	Carbon Emissions	Complaint Regarding Thermal Comfort	Cost Effective Development	Electrical Losses	Electricity From Renewable Sources	GHG Emissions
Outdoor Temperature	Sensor							
Buildings Electricity Demand	Sensor		x		x	x	x	x
Buildings Heat Demand	Sensor		x		x			x
Sources Electricity Produced	Sensor	x	x		x	x		x
Sources Heat Produced	Sensor	x	x		x			x
Renewable Sources Electricity Produced	Sensor						x	
Renewable Sources heat Produced	Sensor							
Biomass Waste Volume Consumed	Sensor							
Buildings Air Radian Temperature	Simulation			x				
Buildings Air Indoor Temperature	Simulation			x				
Buildings Air Indoor Relative Humidity	Simulation			x				
Ebbw Vale Waste Volume Produced	[604]							
Rural Zone Air Outdoor Temperature	Mocked							
National Grid Co2 Emissions	[605], [606]		x					
National Grid GHG Emissions	[605], [606]							x
National Grid Energy Cost (ELEC)	[607]				x			
Energy Sources Co2 Emissions	[605], [606]		x					
Energy Sources GHG Emissions	[605], [606]							x
Energy Sources Energy Cost	[607]				x			
Energy Source NOx Emission	[605], [606]	x						

Data	Provenance	Heating From Renewable Sources	Heat Losses	Onsite Generation	Organic Waste For Energy Generation	Thermal Gradient Differences	Timeout Thermal Comfort Levels	Total Energy Demand
Outdoor Temperature	Sensor					x		
Buildings Electricity Demand	Sensor			x				x
Buildings Heat Demand	Sensor	x	x	x				x
Sources Electricity Produced	Sensor			x				
Sources Heat Produced	Sensor		x	x				
Renewable Sources Electricity Produced	Sensor							
Renewable Sources heat Produced	Sensor	x						
Biomass Waste Volume Consumed	Sensor				x			
Buildings Air Radian Temperature	Simulation							
Buildings Air Indoor Temperature	Simulation						x	
Buildings Air Indoor Relative Humidity	Simulation						x	
Ebbw Vale Waste Volume Produced	[604]				x			
Rural Zone Air Outdoor Temperature	Mocked					x		
National Grid Co2 Emissions	[605], [606]							
National Grid GHG Emissions	[605], [606]							
National Grid Energy Cost (ELEC)	[607]							
Energy Sources Co2 Emissions	[605], [606]							
Energy Sources GHG Emissions	[605], [606]							
Energy Sources Energy Cost	[607]							
Energy Source NOx Emission	[605], [606]							

Consequently, among the 193 original indicators, 17 KPIs, mostly related to energy aspects, have been chosen related to case studies. Equally, rank questions have been used to increase the response rate, avoiding open-ended questions when possible. Additionally, experts were invited individually based on specific publication and work they undertook with a personalised email. One reminder was sent to the expert after approximately a month to renew the research interest in their participation. Efforts have been put on transparency, explicating the use of the data for the thesis purpose. Finally, a more significant amount of experts have been contacted in order to gather a minimum of 30 questionnaires completed.

The consultation will be structure around 2 round of survey:

- the 1st round will proceed to the selection of the most relevant indicators considering their global relevance, measurability and improvability.
- the 2nd round will revise the answer of the first round in order to reach consensus. Additionally, it will look at the indicators previously selected and define more precisely contextual parameters such as the time scale to be considered, forecasting needs, weight etc.

Overall, the study aims to identify and well define the indicators suitable for a dynamic and near-real-time framework that will serve operational and managerial actions based on the criteria shown in Table 4-6.

Table 4-6 DELPHI consultation criteria

DELPHI Round	Criteria	Comment
Round 1	Relevance	Question the overall relevance of the KPI in regards to urban sustainability
	Improvement	Question if the KPI can be improved through managerial and operational efforts.
	Measurability	Question the ability to capture the KPI in real-time
Round 2	Relevance	-
	Improvement	-
	Measurability	-
	Frequency	Question on which time-frequency the KPI should ideally be logged.
	Forecast Horizon	Question the ideal forecast time horizon for operational purposes.

4.4.5 1st Round Survey

The first round survey is intended to question the relevance of the KPIs within a real-time sustainability assessment scheme including their global relevance, measurability and their ability of improvement via operational and managerial interventions.

4.4.5.1 *1st Round Survey Design*

Three questions have been asked to the participants:

Q1 *Do you think these indicators are irrelevant/neutral/relevant/strongly relevant for a real-time sustainability assessment framework at an urban scale?*

The first point of the consultation questions the general relevance of the indicator for sustainability assessment in real-time. Indeed, all indicators are not worst being measured in real-time because of their static nature or the impossibility to catch their state in real-time. Therefore such question aims at putting aside those types of KPIs.

Q2 *Do you think these indicators are hardly measurable/ measurable but costly/ available or easily measurable in real or near real-time?*

Each indicator would ideally be computed from sources that provide real-time data, hence the second question that focuses on the ability to capture the indicator in real-time for a realistic implementation.

Q3 *Do you think these indicators are not improvable/possibly improvable/easily improvable with managerial and operational efforts?*

The third question focus on the degree on which certain indicators can be improved through managerial and operational efforts. Indeed, the framework is intended to be used for operational decision-making and therefore it is not essential to present to the user an indicator on which he cannot influence by this mean.

Those questions are presented in a ranked form where the respondent has to select the extent of relevance, measurability and improvability of each KPI.

Note that for a more understandable consultation, the *Cost Effective Development indicator* has been divided into *Gas, Electricity and Energy Losses Costs* bringing the number of indicators addressed to 19.

4.4.5.2 *1st Round Survey Results*

The first question interrogates on the overall relevance of the indicators for a near real-time implementation. Figure 4-5 shows side by side the frequencies of the responses on the KPIs relevance. The electrical losses indicator gathers the most mitigated input with 4 participants estimating it is irrelevant and 6 neutral for a total of 10 out of 29 answers or around 35%. On the other hand, 100% of the responding experts estimate the share of electricity from renewable sources relevant or strongly relevant. This is followed by *HeatFromRenewableSources, GHGEmitted, AirQualityComplaint, AirQualityLevelOut, CO2Emitted* and *TotalEnergyDemand* that are considered to be relevant by at least 90% of the respondents.

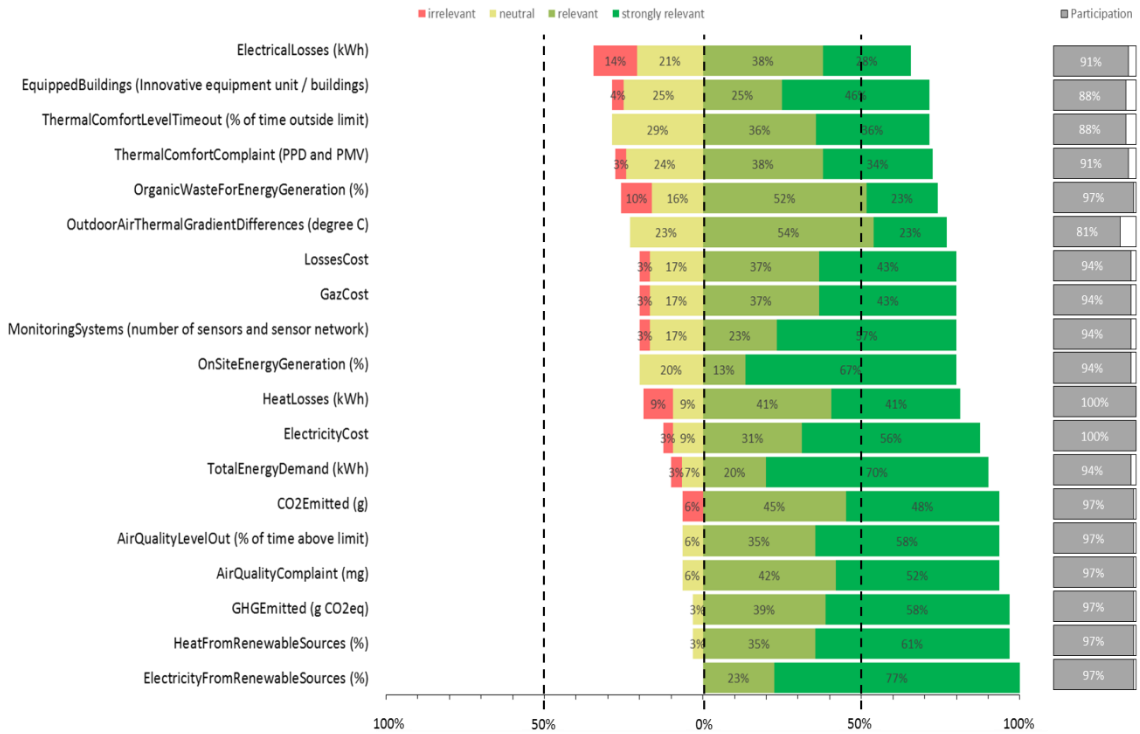
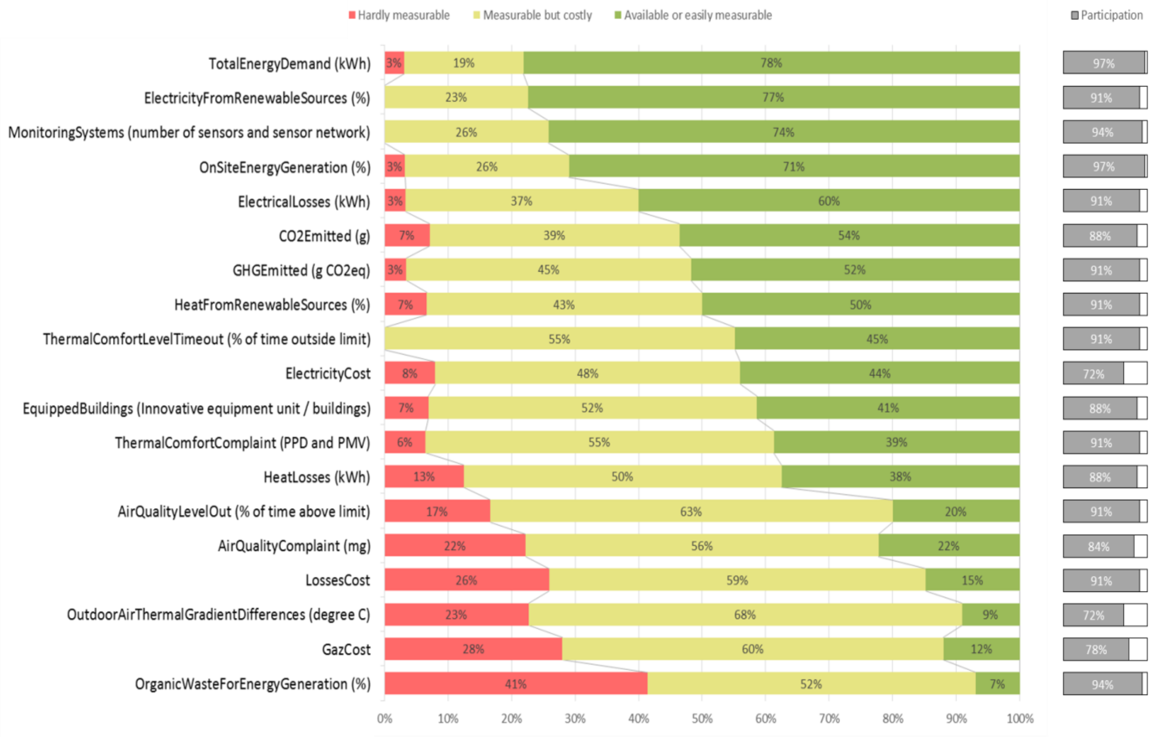


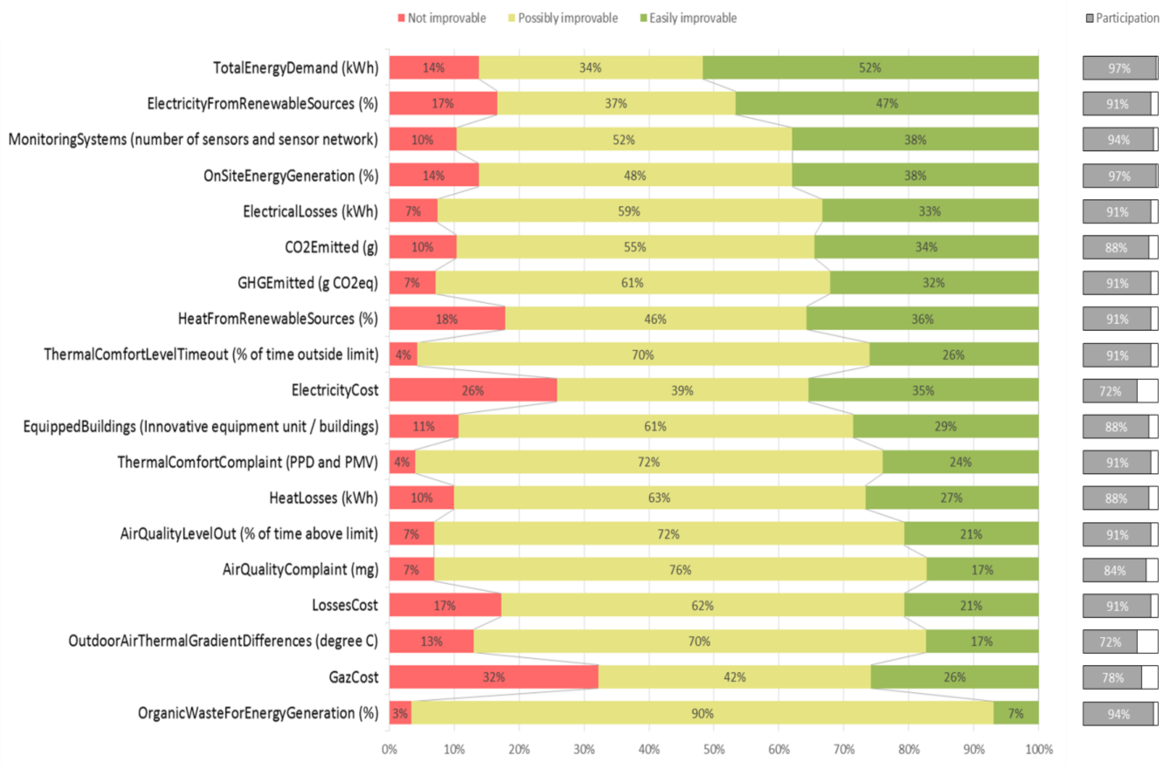
Figure 4-5 Frequency distribution for indicator's relevance

Overall, all indicators are considered strongly relevant by more than 23% (*OutdoorAirThermalGradientDifferences* and *OrganicWasteForEnergyGeneration*) and relevant by more than 66% (*ElectricalLosses*) with an average at 49% and 83% respectively. Note that the meaning of the "Outdoor Air Thermal Gradient" has been questioned by one participant which may explain it has the lowest participation ratio with only 26 responses out of 32 or 81%. Equally, one participant estimated unclear the definition of electrical losses and if it was rather losses in transmission or wasted electricity. Another wondered if the CO2 emissions were focusing on occupants or building systems. Finally, another interesting outcome is that 11 KPIs out of 19 are considered as being both irrelevant by some participants and strongly relevant by some others. This shows that the field of sustainability still does not consensus within the experts' community for the selection of its KPIs.

Figure 4-6 presents the frequencies on the different KPIs measurability and improvement. In term of measurability, energy-related indicators like *ElectricityFromRenewableSources*, *OnSiteEnergyGeneration*, *TotalEnergyDemand* or *HeatFromRenewableSource* as well as costs-centred such as *GazCost* and *ElectricityCost* and air quality related indicators such as *AirQualityComplaint* and *AirQualityLevelOut* are the most easily measurable with at least half (50%) of the respondent considering this option. All the remaining indicators, apart from *OutdoorAirThermalGradientDifferences*, are believed to be measurable but costly by at least 50% of the respondents. *OutdoorAirThermalGradientDifferences* is the only indicator where no majority is outlined. *EquippedBuildings, OrganicWasteForEnergyGeneration, ElectricalLosses, ThermalComfortLevelTimeout, LossesCost, ThermalComfortComplaint* and *HeatLosses* have the most negative figures with 13% to 41% of respondents thinking they are hardly measurable. Note that losses related KPIs are all part of this last list. The other KPIs present less than 8% entries for this option.



(a)



(b)

Figure 4-6 KPIs measurability (a) and improvement (b)

In term of improvability via managerial efforts, only *TotalEnergyDemand* is considered to be easily improvable by more than half of the respondent with 52%. *OrganicWasteForEnergyGeneration, AirQualityComplaint, AirQualityLevelOut, ThermalComfortComplaint, OutdoorAirThermalGradientDifferences, ThermalComfortLevelTimeout, HeatLosses, LossesCost, EquippedBuildings, GHGEmitted, ElectricalLosses, CO2Emitted* and *MonitoringSystems* indicators gather 90% down to 52% of “possibly improvable” responses. The remaining indicators that are *OnSiteEnergyGeneration, HeatFromRenewableSources, GazCost, ElectricityCost* and *ElectricityFromRenewableSources* do not predominantly fit in any categories. The “Not improvable” option ranges from 3% for *OrganicWasteForEnergyGeneration* to 32% for *GazCost* with a median at 10% with a 3rd quartile at 17%. Therefore *ElectricityCost* and *GazCost* are remarkable with 26% and 32% of respondent considering they are not improvable. Those two indicators are actually interesting as they cover all categories in almost equal proportions. This can prefigure a lack of understanding of the question.

Note that for both questions *OutdoorAirThermalGradientDifferences*, *ThermalComfortComplaint*, *ThermalComfortLevelTimeout* and *EquippedBuildings* have the lowest participation ratio with 22 to 25 answers out of 32.

Figure 4-7 aims a better representing the overall outcomes of the study on improvement/measurability by crossing their occurrences. The graph shows that for all the indicators, a minimum of 50% of the experts consider them as at least possibly improvable via managerial actions and measurable but costly in certain cases. *TotalEnergyDemand, HeatFromRenewableSources, ElectricityCost, ElectricityFromRenewableSources* and *OnSiteEnergyGeneration* are the ones with the most confidence with respectively 54%, 54%, 58%, 60% and 66% of the respondents that believe they are at least possibly improvable and easily measurable. *MonitoringSystems, HeatLosses, GHGEmitted, ElectricalLosses, CO2Emitted, ThermalComfortComplaint, LossesCost, OrganicWasteForEnergyGeneration* and *ThermalComfortLevelTimeout* are all believed by a majority of respondents to be at least possibly improvable but that will likely to be costly to measure. It is worth mentioning that *ElectricityCost and GasCost* have really mixed opinions with some of the best and worst results.

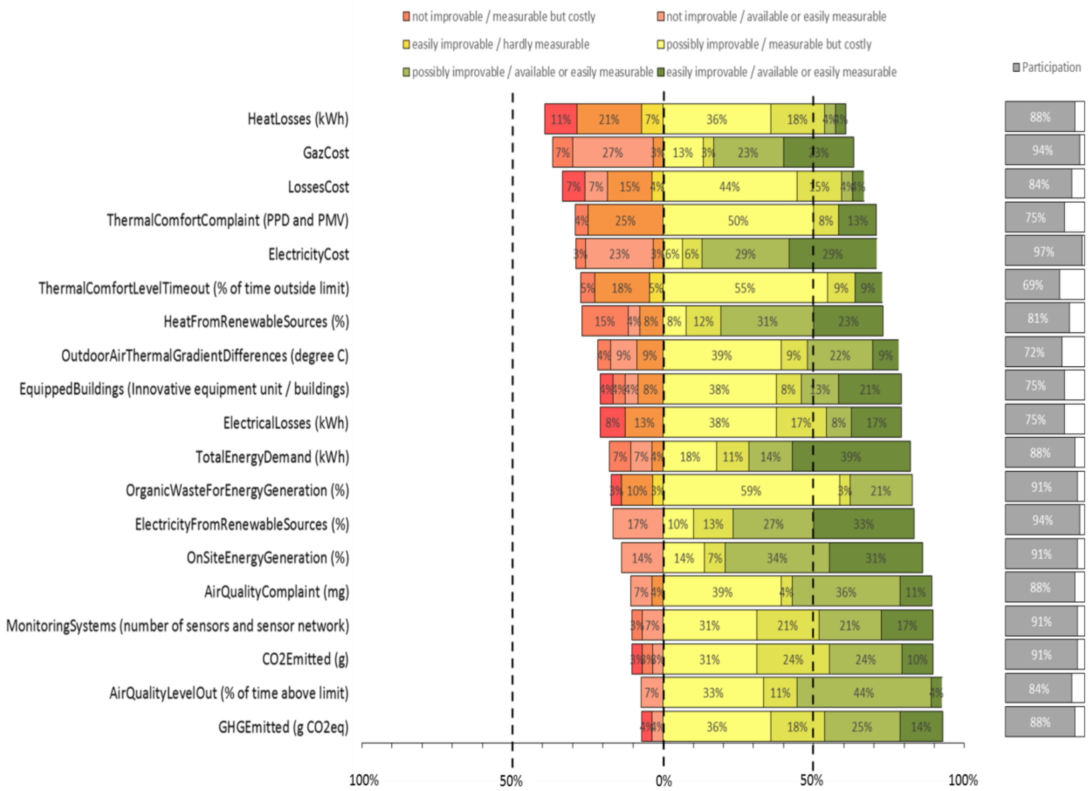


Figure 4-7 Frequency distribution for indicator's cross table improvement/measurability

Additional concerns have been expressed by the experts that could limit the definition of the framework in real or near real-time:

– Privacy concerns often come with the installation of meters which results in a strong resistance of the population against such tools.
– If meters such as energy/electricity/gas/water meters are nowadays relatively common, it remains challenging to deploy them in a non-intrusive way. Therefore, the costs of installation or upgrade of the meters is often left to a third-party instead of energy suppliers.
– Some argue that if the technology is there to accurately measure energy consumption and production, substantial efforts remains for their management, especially in the case of smart grids.
– The meaning of real or near-real-time itself has been questioned, arguing that it is an abstract notion that differs from an indicator to another. For instance, energy demand observations much depend on the data transmission interval between user premises and Power Company.

- The opinions on organic waste for energy generation are mixed, some arguing that the logistic to get data makes it impossible to reach, other giving concrete examples of efforts already implemented in the U.S. or Sweden.
- When it comes to air quality, pollutant emissions and thermal comfort, Governments' regulation may be involved.

Overall, all the indicators presented to the participants have been validated as being relevant for future real-time assessment schemes. Additionally, most of the experts have considered that the indicators were measurable in real-time but that it may be costly; and that they can possibly be improved through operational and managerial interventions. The second round of the survey will be submitted to the same participants in order to clarify some misunderstanding and to deepen the knowledge on the indicators implementation.

4.4.6 2nd Round Survey

The second round of the survey is intended to solve some misunderstanding raised by the participant in the 1rst round and to extend the consultation to additional questions.

4.4.6.1 *2nd Round Survey Design*

In the first survey, some feedback pointed out the fact that some indicators were not well defined:

- *GasCosts* was misleading and is now called *HeatCost* which are the costs induce by heating systems in operation at district level (not all heating systems run on gas).
- *ElectricityCosts* are the costs induce by electricity systems in operation at district level.
- *LossesCosts* are the costs associated with both heat and electricity losses at district level.
- *CO2Emitted* are the emission due to buildings' operation, mainly related to energy consumption (Emissions for transportation are not included in this case).
- Idem for *GHGEmitted*.
- *OutdoorAirThermalGradientDifferences* (degree C) is the temperature difference between an urban area and the nearest rural area. It is meant to reflect the urban heat island effect.

Based on these new definitions, questions 1 to 3 have been asked again.

Moreover, a new set of question has been asked in order to better understand in which extent the KPIs can be considered in real or near-real-time:

Q4 *Real-time is an abstract notion and KPIs are often measured at a certain frequency. In your opinion, from which time-frequency can we reasonably consider the KPI as being relevant and exploitable in real or near-real-time: less than minutely, minutely to hourly, hourly to daily or more than daily?*

And which forecast horizon should be considered:

Q5 *In your opinion, which forecast time horizon is relevant for this indicator in order to perform efficient decisions: less then minutely, minutely to hourly, hourly to daily or more than daily?*

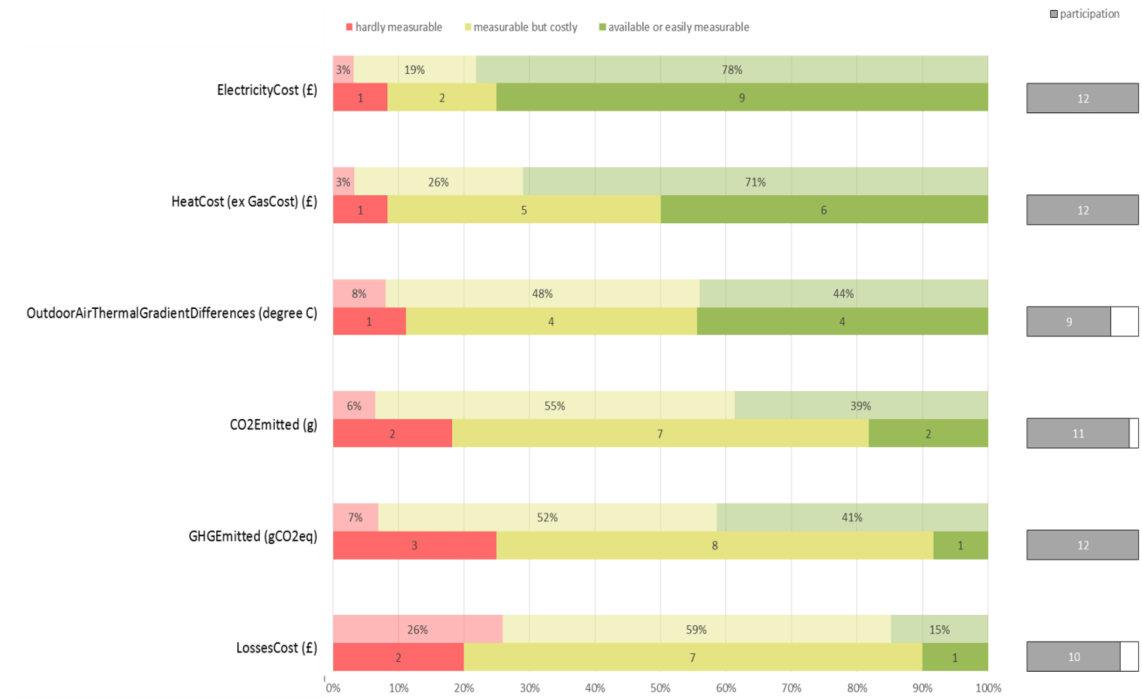
Note that for this second round only 12 experts participated in the survey.

4.4.6.2 *2nd Round Survey Results*

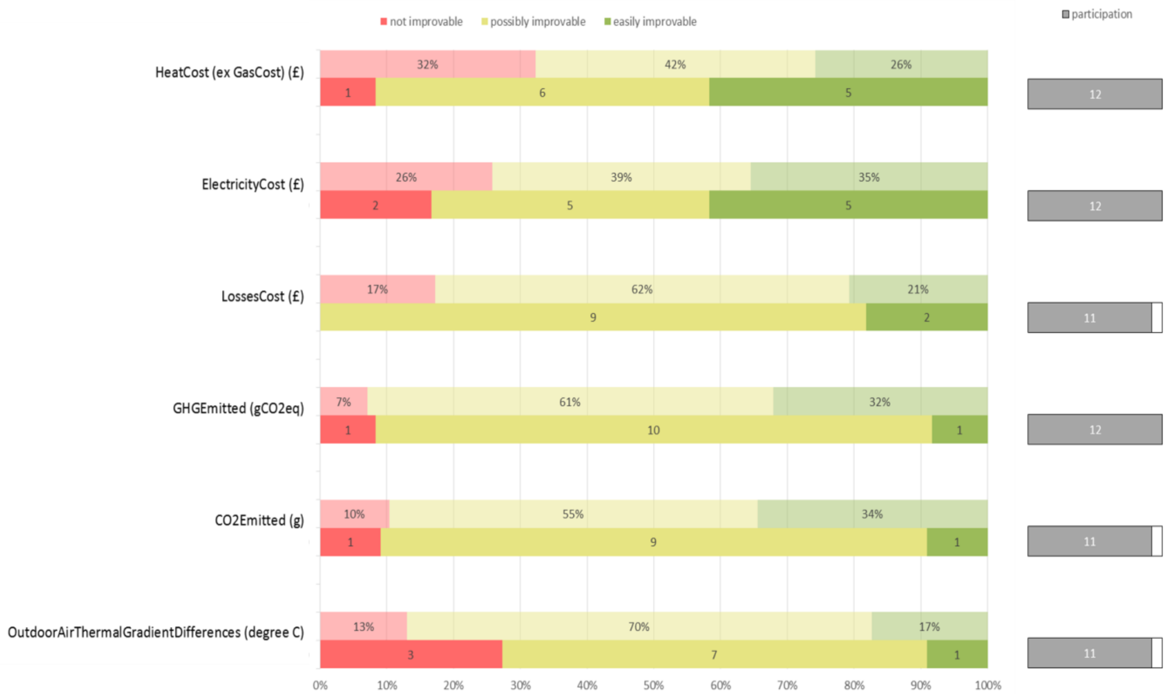
First, a second iteration of the survey has been done for the indicators that were gathering the most mixed reviews namely *OutdoorAirThermalGradientDifferences, CO2Emitted, GHGEmitted, LossesCost, ElectricityCost* and *HeatCost (previously named GasCost).* If they have been largely considered as relevant, their improvability and measurability will be the subject to a second round.

Figure 4-8 shows the degree of measurability and improvability of the previously cited KPIs. After clarification of the KPIs nature, it seems that, in term of measurability, the experts' opinions did not change significantly apart for the emission-related indicators. Indeed, *GHGEmitted* and *CO2Emitted* indicators have seen a greater proportion of responses toward costly measurements and hardly improvable at the expense of easily measurable. On the improvability, *HeatCost* and *ElectricityCost* have more inputs (in proportion) toward easily and possibly improvable, reducing the proportion of not improvable opinions. The "not improvable" responses for *LossesCost* have shifted to "possibly improvable" with 9 occurrences out of 11. Respondents have considered in a greater extent that *GHGEmitted* and *CO2Emitted* indicators were possibly improvable with respectively 10 out of 12 and 9 out of 11 occurrences, at the expense of the "easily improvable" option. Finally, *OutdoorAirThermalGradientDifferences* have lost points on easily improvable for hardly improvable while the “possibly improvable” option remained unchanged.

Overall, the low response rate of this last survey has significantly reduced the analysis relevance. Nevertheless, it seems quite clear that the majority of the experts agree on the fact that all the indicators are measurable in real-time, some for a significant cost, and that they are at least possibly improvable during the operation stage without the intervention of great structural changes.



(a)



(b)

Figure 4-8 Second round KPIs measurability (a) and improvement (b) (1st round vs 2nd round)

The fourth question relates to the notion of real-time and the extent to which a certain KPI can be considered as measured in real-time. Figure 4-9 gives the KPIs log frequencies considered by the expert for each indicator. Apart from *CO2Emitted* where 7 out of 12 experts think that daily to hourly measurement frequency is reasonable and *OrganicWasteForEnergyGeneration* where 5 out of 9 think that more than daily measurement frequency is reasonable, none of the other KPIs demonstrates a clear trend toward an interval.

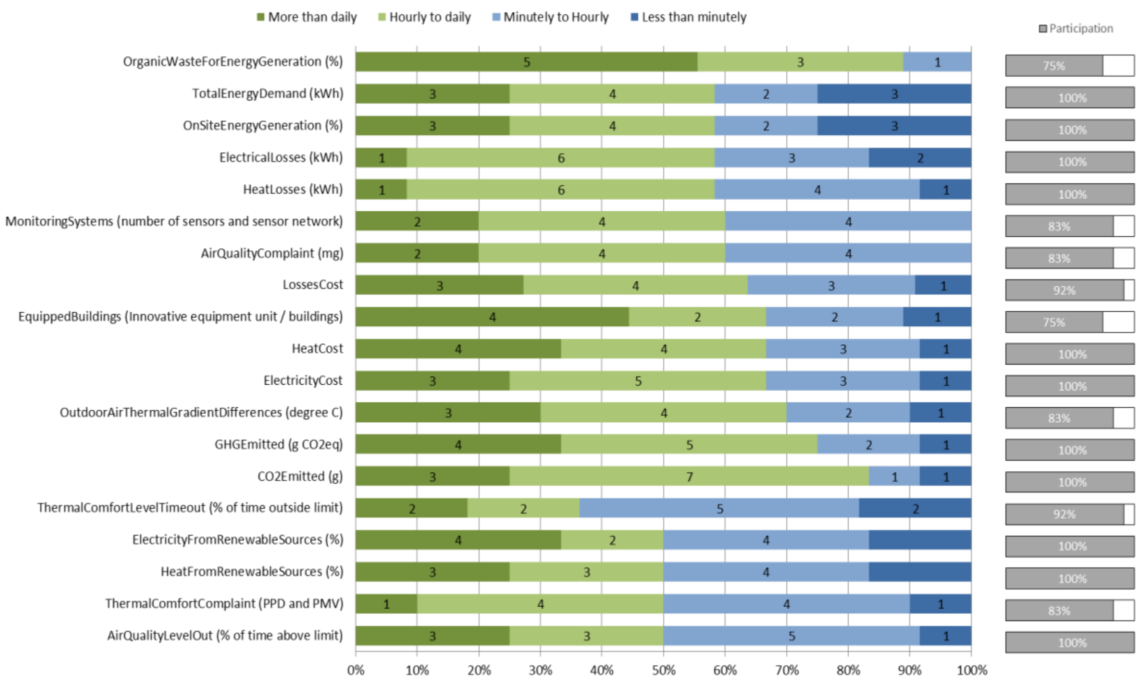


Figure 4-9 KPIs logs frequency

Therefore, the dataset is investigated against the following test:

What is the lowest log frequency that the majority of the experts will consider as being reasonable for the KPI real-time assumption?

Table 4-7 summarises the answers to this question.

Table 4-7 Answer statements for question 4

	At least ___% of the experts are satisfied by____measurements or higher frequency.	
AirQualityLevelOut (% of time above limit)	92%	At least Hourly
ThermalComfortComplaint (PPD and PMV)	90%	At least Hourly
HeatFromRenewableSources (%)	83%	At least Hourly
ElectricityFromRenewableSources (%)	83%	At least Hourly
ThermalComfortLevelTimeout (% of time outside limit)	82%	At least Hourly
CO2Emitted (g)	83%	At least Daily

GHGEmitted (g CO2eq)	75%	At least Daily
OutdoorAirThermalGradientDifferences (degree C)	70%	At least Daily
ElectricityCost	67%	At least Daily
HeatCost	67%	At least Daily
EquippedBuildings (Innovative equipment unit / buildings)	67%	At least Daily
LossesCost	64%	At least Daily
AirQualityComplaint (mg)	60%	At least Daily
MonitoringSystems (number of sensors and sensor network)	60%	At least Daily
HeatLosses (kWh)	58%	At least Daily
ElectricalLosses (kWh)	58%	At least Daily
OnSiteEnergyGeneration (%)	58%	At least Daily
TotalEnergyDemand (kWh)	58%	At least Daily
OrganicWasteForEnergyGeneration (%)	56%	More than Daily (2 experts specified monthly, 1 weekly)

Figure 4-10 highlights the outcome of the descriptive data analysis presented in Table 4-7. Three trends are identifiable, *AirQualityLevelOut, ThermalComfortComplaint, HeatFromRenewableSources, ElectricityFromRenewableSources* and *ThermalComfortLevelTimeout* with at least hourly measurement; *OrganicWasteForEnergyGeneration* with at least more than daily measurement; and the other KPIs with at least daily measurement.

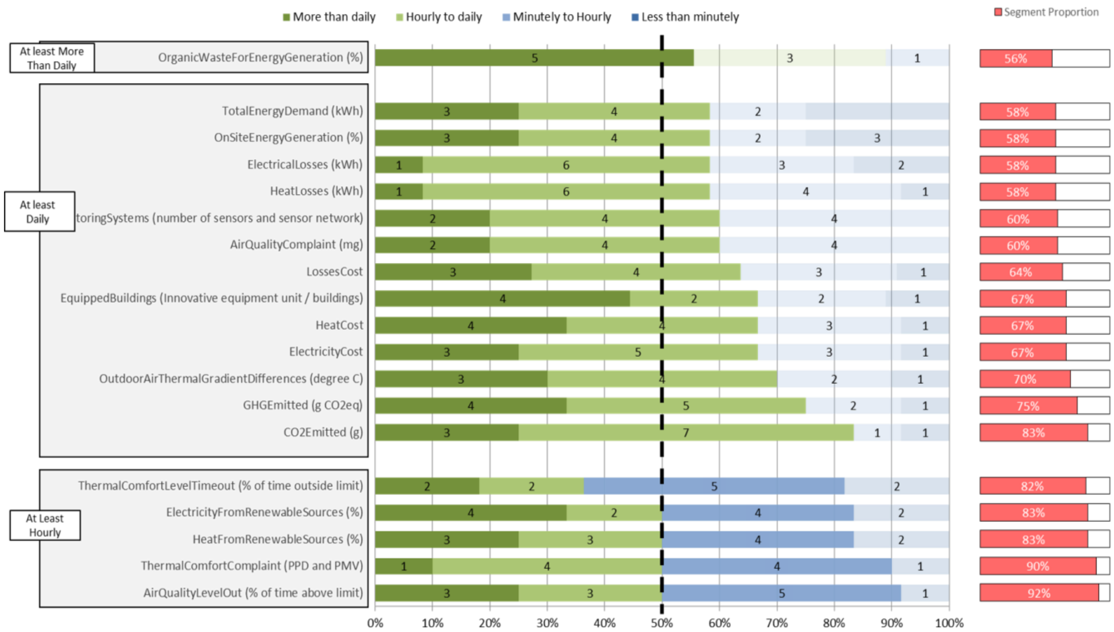


Figure 4-10 KPIs logs frequency bis

Note that in the case of *OrganicWasteForEnergyGeneration* where more than daily measurement predominates, two experts have specified that monthly measurement is the lowest frequency reasonable to consider the KPI as in real-time; one respondent has specified “at least weekly” measurements and the two remaining experts have not specified anything.

Finally, question 5 relates to the time horizon for which a forecast is relevant in order to perform good operation and maintenance of an urban area. Figure 4-11 gives an overview of the expert answers. *GHGEmitted, HeatCost, ElectricityCost, OrganicWasteForEnergyGeneration* and *EquippedBuildings* distinguish themselves by gathering more than half of the opinion toward “more than daily” forecast horizon. The other indicators do not demonstrate any preference toward a specific segment.

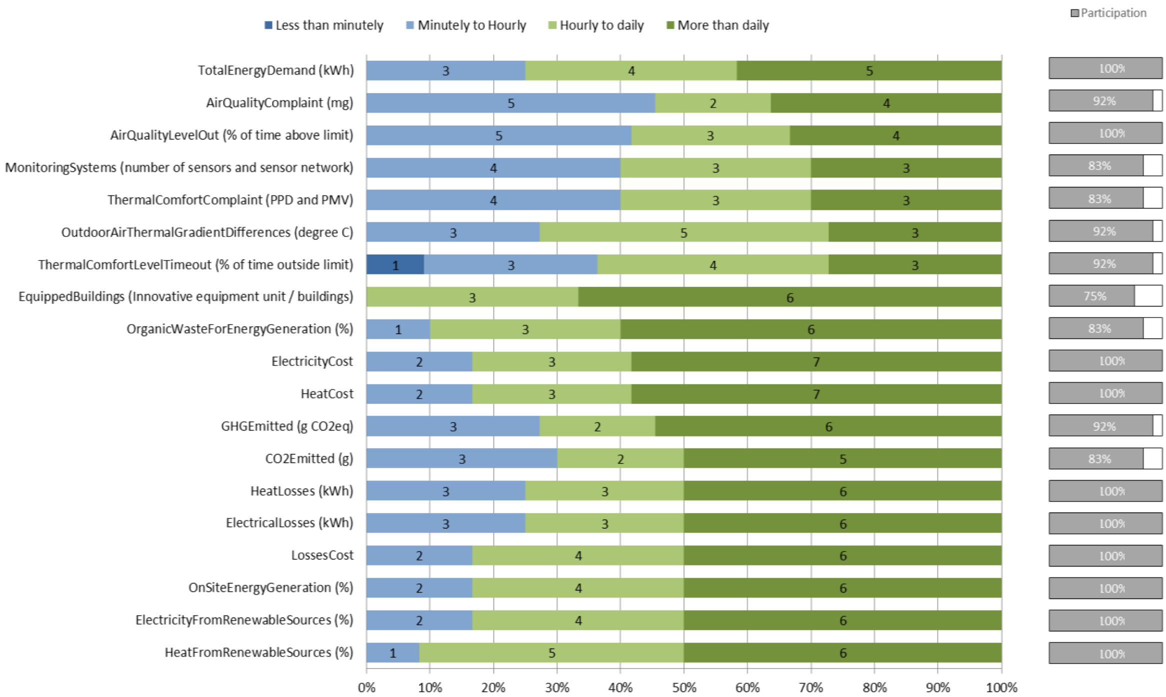


Figure 4-11 KPIs forecast time horizon

Therefore, the dataset is investigated against the following test:

What is the lowest time horizon that the majority of the experts will consider as being reasonable for the KPI forecast time horizon?

Table 4-8 summarises the answers to this question.

Table 4-8 Answer statements for question 5

	At least ___% of the experts are satisfied by___forecast time horizon or higher.	
HeatFromRenewableSources (%)	100%	More than a day
ElectricityFromRenewableSources (%)	100%	More than a day
OnSiteEnergyGeneration (%)	100%	More than a day
LossesCost	100%	More than a day
ElectricalLosses (kWh)	100%	More than a day
HeatLosses (kWh)	100%	More than a day
CO2Emitted (g)	100%	More than a day
GHGEmitted (g CO2eq)	100%	More than a day
HeatCost	100%	More than a day
ElectricityCost	100%	More than a day
OrganicWasteForEnergyGeneration (%)	100%	More than a day
EquippedBuildings (Innovative equipment unit / buildings)	100%	More than a day
ThermalComfortLevelTimeout (% of time outside limit)	73%	At least a day
OutdoorAirThermalGradientDifferences (degree C)	73%	At least a day
ThermalComfortComplaint (PPD and PMV)	70%	At least a day
MonitoringSystems (number of sensors and sensor network)	70%	At least a day
AirQualityLevelOut (% of time above limit)	67%	At least a day
AirQualityComplaint (mg)	64%	At least a day
TotalEnergyDemand (kWh)	58%	At least a day

Figure 4-12 highlights the outcome of the descriptive data analysis presented in Table 4-8. For *ThermalComfortLevelTimeout, OutdoorAirThermalGradientDifferences, ThermalComfortComplaint, MonitoringSystems, AirQualityLevelOut, AirQualityComplaint* and *TotalEnergyDemand,* the majority of the experts consider that forecasts must be at least up to a day.

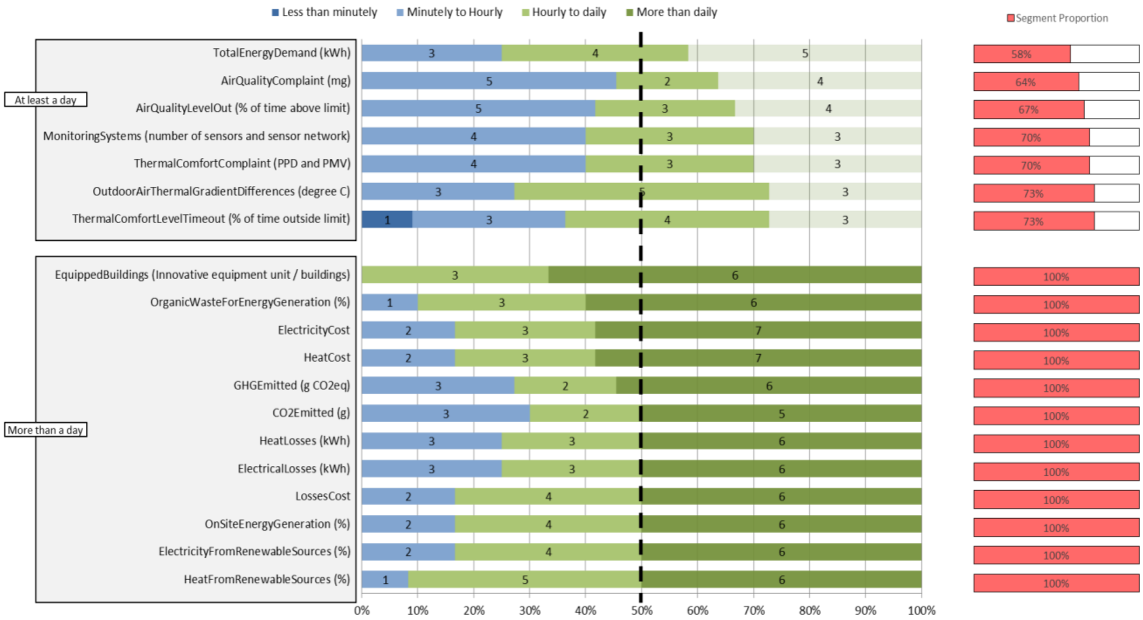


Figure 4-12 KPIs forecast time horizon bis

Because it is quite unclear what the “more than daily” option encompasses, the experts have been asked to precise the time horizon when choosing this option. The answers given range from *“Monthly”*, *“Weekly”*, *“Weekly or more”* or *“Daily or Weekly or Monthly (but not hourly)”*. Figure 4-13 gives the experts opinion of the KPIs forecast horizon with a detailed description of what “more than daily” encompasses. Following this detailed version, *HeatFromRenewableSources, ElectricityFromRenewableSources, OnSiteEnergyGeneration, LossesCost, ElectricalLosses, HeatLosses, CO2Emitted* and *GHGEmitted* should be predicted to at least a day as most of the expert would agree with that. For *HeatCost, ElectricityCost* and *OrganicWasteForEnergyGeneration,* the forecast horizon remains obscure as not all experts have mentioned a precise horizon when selecting “more than daily”. Nevertheless, it is reasonable to say that those should be predicted to at least a day or a week. Same goes for *EquippedBuildings* where it is still unclear if the horizon should be at least a day or a month.

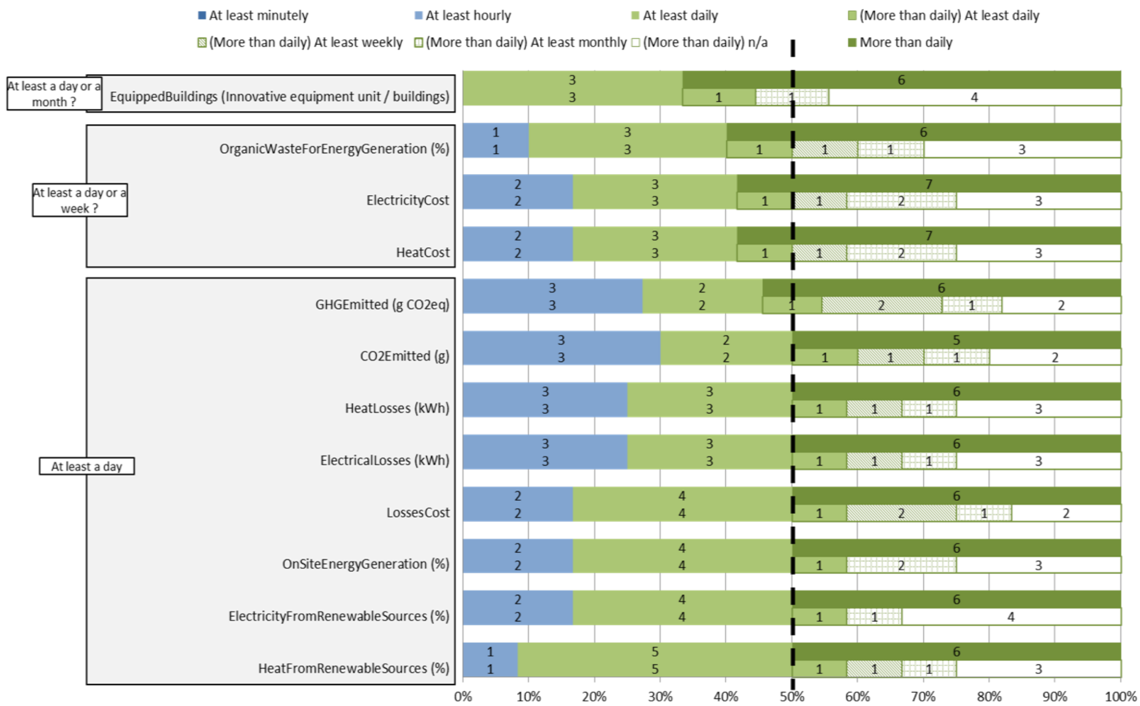


Figure 4-13 KPIs forecast time horizon ters

Table 4-9 summarizes the outcome of question 4 and 5 giving the reasonable time frequencies and forecast time horizon for each of the KPIs.

Table 4-9 Q4 & Q5 outcome summary

	At least ___% of the experts are satisfied by____measurements or higher frequency.		At least ___% of the experts are satisfied by___forecast time horizon or higher.	
AirQualityComplaint (mg)	60%	At least Daily	64%	At least a day
AirQualityLevelOut (% of time above limit)	92%	At least Hourly	67%	At least a day
CO2Emitted (g)	83%	At least Daily	60%	At least a day
ElectricalLosses (kWh)	58%	At least Daily	58%	At least a day
ElectricityCost	67%	At least Daily	100%	More than a day
ElectricityFromRenewableSources (%)	83%	At least Hourly	58%	At least a day
EquippedBuildings (Innovative equipment unit/buildings)	67%	At least Daily	100%	More than a day
GHGEmitted (g CO2eq)	75%	At least Daily	55%	At least a day
HeatCost	67%	At least Daily	100%	More than a day
HeatFromRenewableSources (%)	83%	At least Hourly	58%	At least a day
HeatLosses (kWh)	58%	At least Daily	58%	At least a day
LossesCost	64%	At least Daily	58%	At least a day
MonitoringSystems (number of sensors and sensor network)	60%	At least Daily	70%	At least a day

OnSiteEnergyGeneration (%)	58%	At least Daily	58%	At least a day
OrganicWasteForEnergyGeneration (%)	56%	More than Daily	100%	More than a day
OutdoorAirThermalGradientDifferences (degree C)	70%	At least Daily	73%	At least a day
ThermalComfortComplaint (PPD and PMV)	90%	At least Hourly	70%	At least a day
ThermalComfortLevelTimeout (% of time outside limit)	82%	At least Hourly	73%	At least a day
TotalEnergyDemand (kWh)	58%	At least Daily	58%	At least a day

4.5 SUMMARY

In this chapter, the author has investigated the existing urban sustainability assessment scheme with the objective to develop a new framework. The end goal is the creation of a framework that does consensus, can be implemented in real-time for operation and maintenance purposes. This requires a good understanding of what real-time means for each indicator and how their measurement can help during the operation of urban areas. Moreover, in the context of real-time assessment that lies on IoT and big data, one must investigate possible means for measurement of those indicators in order to demonstrate the feasibility of such scheme.

The new framework presented in this section is an outcome of the systematic literature review of Urban Sustainability Assessment frameworks in section 2.1. The 29 references reviewed have served as a base for the determination of the new USA scheme. A top-down approach has been employed, defining first themes and narrowing down to subthemes, criteria and finally indicators. Those are some of the most frequently encountered aspects within the references which prefigure some kind of agreement on the importance. Once the KPIs identified, the author has proceeded to an extended literature review in order to find existing instances of technological assets that can measure them in real or near real-time. Overall, 56 references have been found that support such approach. Finally, an 4 rounds DELPHI consultation of multi-domains experts have been done in order to better grasp significant features of the KPIs such as their simple relevance in real-time, the ability to be measured in real-time, the minimum log frequency to consider them as in real-time and the most relevant forecasting horizon for good operation.

The review of the 29 frameworks has led to the determination of 193 KPIs spread within 90 criteria, 25 subthemes and 8 themes covering aspects from land uses to health and well-being, economic prosperity, climate, efficient urban design, transportation etc.

The KPIs measurement approaches investigation has determined 5 main technological assets that the scheme can leverage on:

- *Sensors* with 20 references including instances such as Electrical Losses with power and voltage sensors [544], heat Losses with temperature/heat sensors [545], Hazardous Waste Generation with RFID and load cell sensor [152], Transit frequency with GPS [539] or Drinking Water Minimum Quality with turbidity sensor [547];
- *Remote sensing and GIS* with 22 references including instances such as Solar Potential On Roof with aerial imagery and object-specific image recognition [559], Accessible Crosswalks with up-to-date GIS object inventory or Mobile Mapping Systems via LiDAR, optical and positioning sensors [564]–[566], Acreage Of Area Covered By Vegetation On-Site with satellite imagery and Normalized Difference Vegetation Index (NDVI) [571] or GHG Emissions with combined surface stations, aircraft, and satellites [567].
- *Statistics and Open Government Data* with 8 references including Enrolment Rates, Transport Affordability, Low Incomes Population via Government Open data platforms APIs [75], [76] or Healthy Weight, Effective Financial Protection For Health Care with Health care open platform [129], [138], [579].
- *BIM* with 5 references including Materials From Local Sources with material take-off schedule [254], [580] or Development and operation Costs with 5D BIM project management costs [168].
- *Crowdsourcing and data mining* with 8 references including Native And Non-Native Plants with smartphone-wielding citizen scientists inventory [557], Recreational Area Satisfaction via social media feedbacks [582] or Leisure Time Physical Activity via social media data mining [584].

Once the KPIs selected and the means for measurability in real-time investigated, they have been submitted for validation to experts. The first outcome was the difficulty to get substantial responses due to the nature of the survey. After two rounds of design investigation, the survey still did not comply with criteria that make a voluntary survey successful. Therefore the choice to select a particular subset of KPIs has been made. Nineteen indicators were selected following the case studies data available. Those have been submitted for validation to the experts in two rounds; the first investigating their relevance, ability to be measured in real-time and improvability while the second investigating required log frequencies and forecast horizon.

All the KPIs have been considered relevant in a real-time scheme at least 66% of the respondent. The vast majority of the experts have considered that the indicators were measurable in real-time but that certain such as *MonitoringSystems, HeatLosses, GHGEmitted, ElectricalLosses* or *CO2Emitted,* for instance, will likely be costly. In term of improvability via managerial and operational efforts, most of the respondents consider the KPIs as at least possibly improvable. When looking at the log minimum log frequency required to reasonably consider the KPIs as in "real-time", 82% or more of the experts would be satisfied with at least hourly measurement for *AirQualityLevelOut ThermalComfortComplaint, HeatFromRenewableSources, ElectricityFromRenewableSources* and *ThermalComfortLevelTimeout.* For the other KPIs, 52% or more are satisfied by at least daily measurements at the exception of *OrganicWasteForEnergyGeneration* where some considered that weekly to monthly would be enough. Finally, the *GHGEmitted, TotalEnergyDemand, ElectricalLosses, ElectricityFromRenewableSources, HeatFromRenewableSources, HeatLosses, LossesCost, OnSiteEnergyGeneration, CO2Emitted, AirQualityComplaint, AirQualityLevelOut, MonitoringSystems, ThermalComfortComplaint, OutdoorAirThermalGradientDifferences, ThermalComfortLevelTimeout* satisfy 53% to 73% of the experts with at least daily forecasts. On the other hand, *ElectricityCost, EquippedBuildings, HeatCost* and *OrganicWasteForEnergyGeneration* are satisfied with more than daily forecast but present too few feedbacks to conclude to a specific horizon. Table 4-10 presents the overall results of the DELPHI consultation.

Table 4-10 DELPHI consultation final outcomes

	Relevance	Measurability	Improvability	Log frequency	Forecast horizon
AirQualityComplaint (mg)	Strongly Relevant	Measurable but costly	Possibly improvable	At least Daily	At least a day
AirQualityLevelOut (% of time above limit)	Strongly Relevant	Measurable but costly	Possibly improvable	At least Hourly	At least a day
CO2Emitted (g)	Relevant	Available or easily measurable	Possibly improvable	At least Daily	At least a day
ElectricalLosses (kWh)	Strongly Relevant	Measurable but may be costly	Possibly improvable	At least Daily	At least a day
ElectricityCost	Relevant	Available or easily measurable	At least Possibly improvable	At least Daily	More than a day
ElectricityFromRenewableSources (%)	Strongly Relevant	Available or easily measurable	At least Possibly improvable	At least Hourly	At least a day
EquippedBuildings (Innovative	Relevant	Measurable but	Possibly	At least	More

equipment unit/buildings)		costly	improvable	Daily	than a day
GHGEmitted (g CO2eq)	Strongly Relevant	Available or easily measurable	Possibly improvable	At least Daily	At least a day
HeatCost	Relevant	Measurable but costly	At least Possibly improvable	At least Daily	More than a day
HeatFromRenewableSources (%)	Strongly Relevant	Measurable but may be costly	At least Possibly improvable	At least Hourly	At least a day
HeatLosses (kWh)	Relevant	Measurable but may be costly	Possibly improvable	At least Daily	At least a day
LossesCost	Relevant	Measurable but costly	Possibly improvable	At least Daily	At least a day
MonitoringSystems (number of sensors and sensor network)	Strongly Relevant	Available or easily measurable	Possibly improvable	At least Daily	At least a day
OnSiteEnergyGeneration (%)	Strongly Relevant	Available or easily measurable	At least Possibly improvable	At least Daily	At least a day
OrganicWasteForEnergyGeneration (%)	Relevant	Measurable but costly	Possibly improvable	More than Daily	More than a day
OutdoorAirThermalGradientDifferences (degree C)	Relevant	Measurable but costly	Possibly improvable	At least Daily	At least a day
ThermalComfortComplaint (PPD and PMV)	Relevant	Measurable but costly	Possibly improvable	At least Hourly	At least a day
ThermalComfortLevelTimeout (% of time outside limit)	Relevant	Measurable but costly	Possibly improvable	At least Hourly	At least a day
TotalEnergyDemand (kWh)	Strongly Relevant	Available or easily measurable	Easily improvable	At least Daily	At least a day

It is unfortunate that only 17 out of the 193 KPIs have been investigated via this survey. Future investigations should be carried out on the model of this survey for the remaining KPIs. The author advocates that such survey be part of a structured organisation or around a cooperative project where participants would be required to answer. This would overcome limitations due to the volunteering nature of the survey and lead to a greater number of exploitable responses over each indicator.

Overall, the chapter has provided evidences to answer the following research questions:

RQ1. What are the issues that face the different stakeholders of an urban system toward sustainability assessment?

The KPI selection based on the extended literature review has highlighted some issues inherent to urban sustainability assessment such as the lack of consensus on the indicators to addressed, unbalanced consideration of sustainability dimension, lack of dynamism, design related perspective, low citizen participation and lack of transparency. The newly developed scheme attempts to bring more consensus and a better coverage of sustainability dimension by synthesising the existing instances. Moreover, the DELPHI consultation strengthen agreement on the KPI selected (at least partially), requiring expert validation. Additionally, open question enabled expert to give their opinion on further matters such as privacy concerns.

RQ2. How an effective urban sustainability assessment can help different parties of a city in their decision making?

The literature review has provided evidences that the KPIs considered in this research are measurable in real time. The real time dimension of the assessment opens new opportunities to hep decision maker. Especially for neighbourhoods in operation where day-to-day operation and maintenance requires an intensive collect of information. Moreover, questions of measurability in real time, improvability via O&M, measurement frequency and forecasting horizon for efficient decision support has been addressed in the DELPHI consultation. Expert opinions enabled to consolidate practical aspects required to efficiently help decision making using urban sustainability assessment.

RQ3. How can sustainability assessment leverage the smart city paradigm, specifically ICTs and the IoT?

The chapter has listed technological means included in the smart city paradigm such as the IoT, GIS technologies, BIM technologies, open government data, Crowdsourcing and data mining. Those have been proven to be valuable means to collect information that can support the determination of sustainability KPIs. Indeed, numerous case studies taken from the literature demonstrate the feasibility to use such technologies for real time measurement of such KPIs.

The next chapter will open on the technological approach for the creation of an USA framework that is able to deal with data heterogeneity. Indeed, this chapter has seen the creation of the USA scheme around several domains and means of measurement. This diversity and heterogeneity will most likely be challenging when developing a common framework that unified those aspects and principle. Therefore, the web semantic approach is investigated in the next chapter with the objective to create a semantic model aligned with the USA scheme that could give meaning to the information and unified the information flow.

Ontological Approach for a Real-Time Urban Sustainability Assessment Research

Chapter 2 has established the importance of interoperability between systems and data sources for the good operation of digital services in the smart city paradigm. Moreover, sustainability covers various domains which will likely involve multiple organisations. Finding ways to unify heterogeneous sources is therefore essential in the future prospect of a real-time sustainability assessment tool. In this chapter, the semantic approach considered to deal with this issue is detailed. The engineering of the USA ontology will be described following the NeOn methodology presented in Section 3.4.6.

The following research questions will be addressed in this chapter:

RQ2. How an effective urban sustainability assessment can help different parties of a city in their decision making?

RQ3. How can sustainability assessment leverage the smart city paradigm, specifically ICTs and the IoT?

RQ4. How can semantic web technologies unify heterogeneous data resources for holistic services and applications?

5.1 OVERALL DESCRIPTION

The smart city along with the provision of information from the IoT, information systems such as BIM or GIS and data mining from the Internet have proven to be a key component to support the creation of a real-time urban sustainability assessment tool that would efficiently help the operation of an urban area as demonstrated in Chapter 2. If such technologies are believed to greatly enhance the discovery of insightful information, they bring technical challenges that need to be overcome. Indeed, in this new prospect, the diversity of data sources and information modelling could limit decision support potential [133], [608]. The question of data heterogeneity, interpretability and transmission is essential [29], [609]–[611] for a seamless data integration into digital services. There is a need for methods and data models that can guarantee efficient interoperability across platforms, domains and scales [131], [157], semantically uniform information flows between organisations [612] and can ensure the cloud-based services quality [613].

Semantic modelling and ontologies therefore come as valuable solutions to overcome such challenges. They define a common semantic data model for specific domains that can explicitly describe real-world concepts and relationships. Because the information is considered in term of its semantic, heterogeneous sources can be aligned over common concepts, improving interoperability. Additionally, such approach will greatly help information discovery as the user does no longer need to know the data structure but its semantic. [614]. Many studies presented in Section 2.3.4 have already investigated the semantic modelling approach of certain sustainability domains such as the *Global City Indicator Foundation* Ontology [364], [365], the *ee-district ontology* [368] or *WISDOM* [27] for instances. The following sections will present the USA ontology, a semantic model to support real-time urban sustainability assessment.

5.2 THE USA ONTOLOGY REQUIREMENT SPECIFICATION

The USA ontology has been developed following the NeOn methodology described in Section 3.4.6. The initial phase in the NeOn methodology is the Ontology Requirement Specification (ORS) [381]. It is a crucial step that helps to define the boundaries of a domain semantic modelling and brings great focus on relevant knowledge resources. It requires the identification of (1) the purpose of the ontology to be developed; (2) the intended uses and users of the ontology; (3) the set of requirements that the ontology should satisfy [524].

5.2.1 Goals, domains and scopes

The ontology will aim to give meaning to diverse components of the urban environment from building to urban furniture, people and processes across multiple domains and scales. It should capture their relationships with sustainability KPIs and how they can be impacted. Overall, it will help in the creation of a holistic model rather than considering aspects individually.

The primary objective of the USA ontology is the semantic representation of the USA framework presented in Chapter 4. All 193 KPIs along with the criteria, subthemes and themes developed should be represented in the ontology. The main task here is to describe semantically which KPIs fits in which criteria, and in turn, which criteria fits in which subthemes and which subthemes fits in which themes. Additionally, each KPIs must be associated with benchmarks that will set sustainability goals. They must be coupled with information about their location and time. The idea is to understand what has been measured, when, where and how it has been measured. On this last matter, sensors networks and ICTs are considered and therefore must be included in the ontology. To summarise, the KPIs and all contextual aspects must be addressed during the development of the ontology.

5.2.2 Use and Users

The primary use here is the creation of a USA tool that would work upon the ontology and its instances to help in sustainability issues discovery and decision making. For that, the user must be able to query the USA scheme itself to better understand how indicators, criteria, subthemes and themes are structured. This would enable an overall understanding of sustainability crucial aspects and their interrelationships. It would promote sustainability as a holistic, interconnected concept. Additionally, the user must be able to query KPIs values and scores for particular locations and times and to compare them with referenced values taken from local policies for instance, or with values taken from different places or times. The discovery of sustainability performance is the most evident use for an assessment scheme and the ontology must facilitate such practice with the use of semantically enhanced information. Finally, an interesting use would be to link the KPIs to actual city objects such as buildings, people and processes, so that the end user can foresee possible interventions on the urban environment that could influence the sustainability performance of the urban area. In this perspective, the user would not only be able to sense performance but also to understand the real world elements that are worth acting upon in order to improve the neighbourhood quality.

Current urban sustainability assessments are mostly designed for the actor of the construction industry that wants to implement good elements and practices in their development. In the current context of a real-time assessment during the operational phase of development, the target user shifts from that vision. Indeed, as Section 2.5 demonstrates, a real-time assessment is closer to a decision support tool than a certification scheme. As such, the USA framework and its ontology users would most likely be district managers that seek to manage the different aspects of a neighbourhood in the most efficient and environmentally friendly way. Sustainability spanning over various domains, this management would be handled by multiple stakeholders such as energy, water, telecom providers as well as community managers, waste collectors etc; in other words, people and organisations in charge of services and operation of a certain urban area will certainly use the ontology. Additionally, citizens could benefit from the provision of such ontology in the simple discovery of their neighbourhood performance. In that particular case, the idea is to raise awareness on sensitive issues at a local scale in order to potentially trigger citizens' behavioural changes and make them reflect upon their relationship with their environment. Overall, the assessment is intended to be used by any actor who participates actively in the city life, from managers and governments to citizens.

5.2.3 Competency Questions

Competency questions are now recognised as a valuable methodological procedure in the engineering of ontologies and have been greatly used in the literature [373], [381], [521], [524], [525]. Section 3.4.6.1 has introduced competency questions as simple questions that the ontology should be able to answer when queried. Broad questions are first introduced and then iteratively refined to draw fundamental questions [525].

The first stage in the development of competency questions has been the creation of a mind map that includes the different elements constituting the intended ontology. Presented in Figure 5-1, the mind map is a conceptual draft; and as such is not intended to reflect the entire complexity of knowledge but rather to help in building a vocabulary repository and in formulating the competency questions. Essential components and their relationships have been roughly formalised such as the link between the scorable elements, which are all elements of the USA scheme that can hold a score (e.g. indicator, criteria, subtheme and theme), their values and benchmarks, the urban objects, possible actions' impacts, etc.

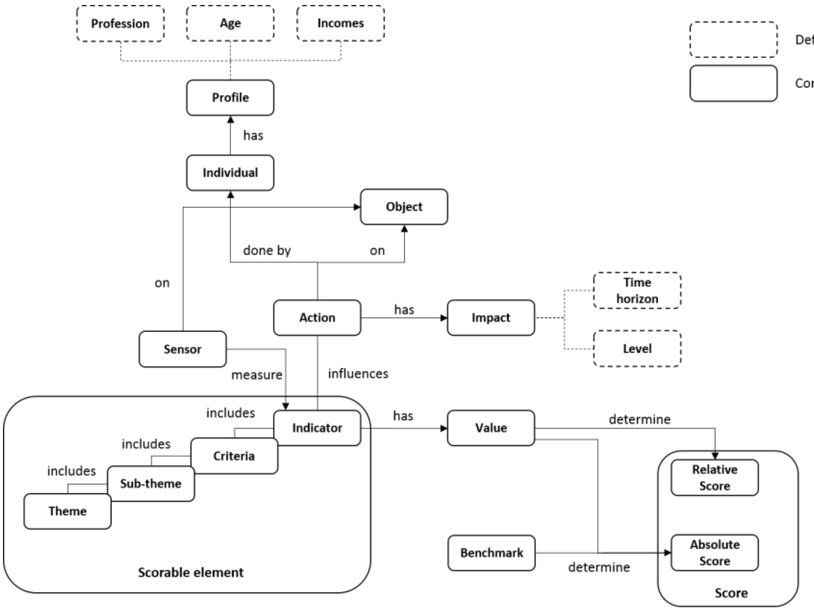


Figure 5-1 Knowledge mind map

Those aspects, aligned with the uses described in the previous section, constitute the base of the competency questions definition. The competency questions have been grouped into three different categories: Table 5-1 relates to the interrelationship between the different scorable elements (Themes, Sub-Themes, Criteria, Indicators) and the discovery of the assessment scheme structure; Table 5-2 presents questions on the scorable elements values and scores and their location, time measured, benchmark and unit; Table 5-3 relates to the link between the scorable elements and various components of the urban environment such as people, objects, actions and their impact on the sustainability performance.

Table 5-1 Scorable elements competency questions

Question	Subject	Property	Object
Which scorable element has scorable element X?	Scorable Element	hasScorableElement	Scorable Element
Which Theme has scorable element X?	Theme	hasScorableElement	Scorable Element
Which SubTheme has scorable element X?	SubTheme	hasScorableElement	Scorable Element
Which Criteria has scorable element X?	Criteria	hasScorableElement	Scorable Element
Which scorable element has criteria X?	Scorable Element	hasCriteria	Criteria
Which Theme has criteria X?	Theme	hasCriteria	Criteria
Which SubTheme has criteria X?	SubTheme	hasCriteria	Criteria
Which scorable element has indicator X?	Scorable Element	hasIndicator	Indicator
Which criteria have indicator X?	Criteria	hasIndicator	Indicator
Which Theme has indicator X?	Theme	hasIndicator	Indicator

Which SubTheme has indicator X?	SubTheme	hasIndicator	Indicator
Which scorable element has SubTheme X?	Scorable Element	hasSubTheme	SubTheme
Which Theme has SubTheme X?	Theme	hasSubTheme	SubTheme

Table 5-2 Value and score competency questions

Question	Subject	Property	Object
Which Indicator has element value X?	Indicator	hasElementValue	ElementValue
Which benchmark is associated with indicator X?	Benchmark	has Associated Indicator	Indicator
Which scorable element has score X?	Scorable Element	hasScore	Score
Which scorable element has absolute score X?	Scorable Element	hasAbsoluteScore	AbsoluteScore
Which scorable element has RelativeScore X?	Scorable Element	hasRelativeScore	RelativeScore
Which scorable element has TemporalRelativeScore X?	Scorable Element	hasTemporalRelativeScore	TemporalRelative Score
Which scorable element has SpatialRelativeScore X?	Scorable Element	hasSpatialRelative Score	SpatialRelative Score
What is the value of Element value X?	Element Value	hasValue	Float/double
What is the Unit of Element value X?	Element Value	hasUnit	String
What is the reference date of Element value X?	Element Value	hasReferenceDate	DateTime
What is the location of Element value X?	Element Value	hasLocation	String
What is the value of Benchmark X?	Benchmark	hasValue	Float/double
What is the Unit of Benchmark X?	Benchmark	hasUnit	String
What is the reference date of Benchmark X?	Benchmark	hasReferenceDate	DateTime
What is the location of Benchmark X?	Benchmark	hasLocation	String
What is the value of Score X?	Score	hasValue	Float
What is the Unit of Score X?	Score	hasUnit	String
What is the reference date of Score X?	Score	hasReferenceDate	DateTime
What is the location of Score X?	Score	hasLocation	String

Table 5-3 Action/Impact and Urban Objects competency questions

Question	Subject	Property	Object
Which action has influenced scorable element X?	Action	hasInfluenceOnScorableElement	ScorableElement
Which action has impact X?	Action	hasImpact	Impact
Which action has associated urban object X?	Action	hasAssociatedUrbanObject	UrbanObject
Which Indicator is associated with urban object X?	Urban Object	hasAssociatedUrbanObject	UrbanObject
Which impact has associated scorable element X?	Impact	hasAssociatedScorableElement	Scorable Element

Which person has done action X?	Person	hasDoneAction	Action
What are the details of Person X?	Person	hasDetails	String
What is the start date of action X?	Action	hasStartDate	DateTime
What is the end date of action X?	Action	hasEndDate	DateTime
What is the level of impact X?	Impact	hasLevel	Float
What is the StartDate of impact X?	Impact	hasStartDate	DateTime
What is the EndDate of impact X?	Impact	hasEndDate	DateTime

Those competency questions have allowed the definition of terminologies for classes and properties that ultimately will help in discovering potentially useful resources for ontology development. In addition, the scope and requirements are increasingly refined with the introduction of such questions which will greatly improve the tonology quality.

5.2.4 Resources Reuse

The NeOn methodology advocates the reuse of existing ontological and non-ontological resources when developing new domain ontology. This will ensure the resulting semantic model is grounded with referenced ontologies and that it is compliant with other domain-specific ontological resources (see Section 3.4.6.3).

Listing the possible uses and users and drawing competency questions has allowed refining the research of resources that can be reused. The different concepts addressed for the future ontology can be categorised. The central element being the representation of the USA scheme structure, it must be linked to mean of measurement and their observations, elements to localise them in space and time and real-world objects linkage.

5.2.4.1 *Urban sustainability resources*

Section 2.3.4 has already listed valuable ontologies for the semantic representation of urban sustainability elements such as *Global City Indicator Foundation* Ontology [364], [365], the *ee-district ontology* [368], *WISDOM* [27], *SEMANCO ontology* [367] or *Transport Disruption Ontology* [370]. Even though those ontologies are valuable options to start semantic modelling they are often restricted to a particular sub-domain of sustainability. Moreover, those are not aligned with the current urban sustainability assessment schemes found in the literature. Consequently, this aspect of the ontology must be carried out from scratch leveraging only on the framework defined in Chapter 4. Future possible alignments with other sustainability-related ontologies may be investigated later on in order to establish the ontology in the domain semantic landscape. All 8 Themes, 26 sub-themes, 90 criteria and 193 indicators and their relationships will be integrated in the USA ontology.

Additionally, benchmarks should be represented and linked to the KPIs. When creating this part of the ontology, it is often useful to align its concepts to upper-level ontologies such as DOLCE+DnS Ultralite ontology (or DUL) [324], [325].

5.2.4.2 *Sensing resources*

The definition of KPIs will rely on sensing technologies and their readings. An essential part of the ontology is the semantic representation of a sensor, its observations and the entire sensing process. Whether it is a direct reading or via calculation, the user must be able to catch the information provenance and flow. As mentioned in Section 2.3.3.1, ontologies such as the Observation and Measurements ontology (O&M) [321], [615], the SSN/SOSA ontology [322] or the SAREF ontology [328] are recognised models for the semantic modelling of observations and sensors.

The O&M ontology, however, is limited as it does not encompass sensor networks and devices or sensing processes. Its prime aim is the modelling of "*observations, and of features involved in sampling when making observations*" [616]. In the case of SAREF, the model describes concrete devices and smart appliances within the built environment [330] and the properties they monitor. Even though the model integrates an exhaustive list of smart appliances and features, its conceptual approach is too grounded into factual instances which can limit its use when unknown characteristics come into play. The SSN ontology appears to be the most valuable option for the development of the USA ontology as it integrates the description of sensors and its observations in a higher degree of abstraction, which in turn enables a more flexible modelling. For instance in SSN, a sensor can be any entity that senses a phenomenon from an individual to a metering device or a computer program.

Figure 5-2 shows a sample of the most interesting classes developed in SSN for the creation of the USA ontology [323].

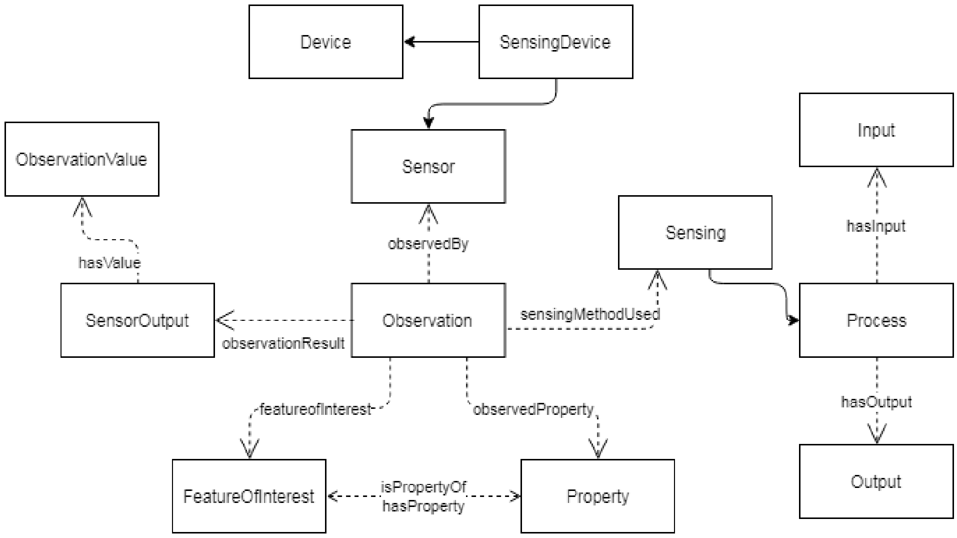


Figure 5-2 The SSN ontology, key concepts and relations adapted from [323]

Like O&M, the triple *Observation*, *Property* and *FeatureOfInterest* is at the core of the schema giving a formal representation of a single capture of an entity feature. A *SensorOuput* results from the Observation and is associated with an *ObservationValue*. The *Observation* is done by a *Sensor* that can be of type *Device* and is the product of a *SensingMethod* which is defined by its *Input* and *Output.* Additionally, sensors can as well be grouped into a system hosted by a common platform and following a certain deployment.

It is worth mentioning that the SSN ontology has been updated in October 2017 [326] after two years of development based on the 2011 version [322]. The USA ontology is therefore grounded in the 2011 version but alignments between the two versions have been thought while developing the updated one.

Finally, the measured values are often associated with a unit of measurement. Those have already been semantically modelled by the NASA with the Quantities, Units, Dimensions and Data Types or QUDT ontology [371].

5.2.4.3 *Spatio-temporal resources*

One of the key requirements specified in the ORS is the ability to geolocate any entity and KPIs. This requires the introduction of geospatial references that could describe geographical information and their relationships. Those features have been the subject of many studies and therefore benefit from well-established models such as GeoSPARQL, as mentioned in Section 2.3.3.3 [353] and depicted in Figure 5-3.

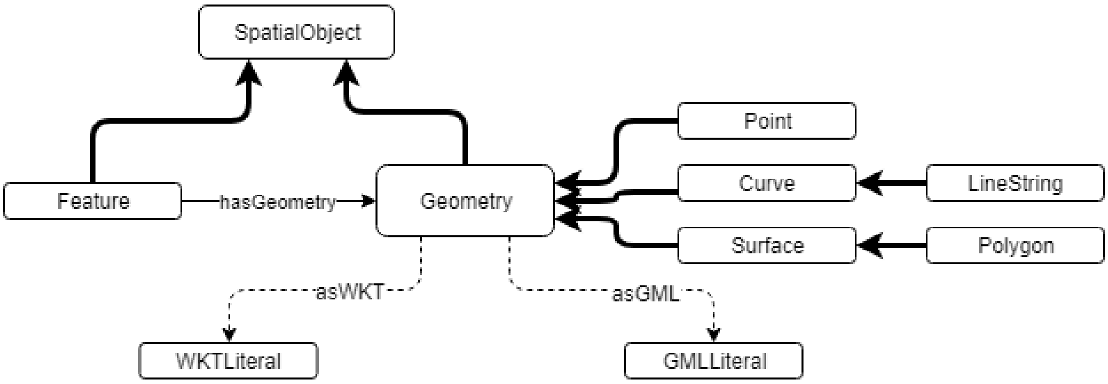


Figure 5-3 GeoSPARQL schema

The semantic model includes geometrical concepts such as point and polygonal areas, as well as spatial relationships such as intersection, union or disjunction between entities. It also integrates datatypes such as Well Known Text (WKT) and Geography Mark-up Language (GML) for the syntactic definition of geometries.

About time-related information, the SSN and DUL ontologies already introduced some key concepts. SSN has introduced object properties such as *ssn:observationResultTime, ssn:observationSamplingTime, ssn:startTime* or *ssn:endTime* that allow the user to point to time-related classes such as *dul:TimeInterval*. The *dul:TimeInterval* itself can be described via the data property *dul:hasIntervalDate* to the datatype *xsd:date.* Those component are believed to be sufficient for the description of time relating to an observation.

Note that those can be aligned with or linked to the Time ontology, *"an ontology of temporal concepts, for describing the temporal properties of resources in the world"* [327]. It includes additional concepts such as hours, days, month, year, time zones, intervals overlaps, intervals disjunction etc.

5.2.4.4 *Urban Objects resources*

One last requirement is the representation of people and objects within the urban environment and their possible connection to the KPIs. Previously described sensors can, therefore, be located in specific objects and the feature of interest that they observe can be detailed and identified as one of those objects. "Urban Objects" includes different elements from the buildings themselves and their interiors to the urban areas' functions and furniture and the people that interact with this environment.

The buildings and their components have already been semantically modelled within the ifcOWL ontology [335], as previously cited in Section 2.3.3.2. IfcOWL is the RDF translation of the Industry Foundation Classes (IFC) standard, a "*data schema and an exchange file format for BIM data*"[617]. The ifcOWL ontology is a large ontology with 1230 classes, 1578 object properties, 5 data properties for the creation of 21,306 axioms and 13,649 logical axioms. This includes, for the most part, the description of the building components geometries and lists of Cartesian points, polylines and so forth [344]. A modular version of the ifcOWL has been created in order to simplify or even remove the geometry for the implementation of a more scalable version [344], [618]. A non-exhaustive list of building components includes classes such as *IfcBeam, IfcDoor, IfcWindow, IfcWall, IfcRoof, IfcOccupant, IfcMaterials, IfcBuilding, IfcBuildingStorey, IfcSpace, IfcFurniture* etc. Overall, all elements of the building environment from component geometry to people and processes and how they relate to each other can be fully semantically described by the ifcOWL ontology. In the present case, it can be useful to locate a *ssn:Sensors* in a specific *ifcOWL:IfcSpace* for instance or *ifcOWL:ifcBuilding* being the *ssn:FeatureOfInterest* of a *ssn:Observation.*

The cityGML ontology presented in Section 2.3.3.3 is believed to help when it comes to model the urban environment. The cityGML ontology is the RDF translation of the cityGML data format, an extension of the GML data format for 3D GIS [357]. The ontology comprises:

- GML based representation of 3D geometries with classes such as *gml:Geometry, gml:MultipleSurface, gml:MultipleCurve, gml:MultiplePoint* etc*;*
- object surface characteristics with *app:Color, app:TextureType, app:Appearance* etc*;*
- terrain model with *dem:ReliefFeature, dem:BreaklineRelief, dem:Elevation* etc;
- sites with classes such as *bdlg:Building, brdg:Bridge* or *tun:Tunnel* for instance*;*
- vegetation with *veg:VegetationObject, veg:PlantCover* and *veg:SolitaryVegetationObject;*
- water bodies *with wtr:WaterBodies, wtr:WaterSurface, wtr:WaterBoundarySurface* etc;
- Transportation facilities with *tran:TransportationObject, tran:TransportationComplex, tran:TrafficArea* or *tran:Railway* for instances;

- City furniture and Generic city objects and attributes with respectively the *frn:CityFurniture* and *gen:GenericCityObject* classes;
- Object grouping with the *grp:CityObjectGroup*.

Additionally, the different sites can be defined by 4 different levels of detail (LOD) with the inclusion of refined elements. For instance, the *bldg:Opening* contains the *bldg:Door* and *bldg:Window* classes, all linked to a *bldg:Building* by *bldg:lod3ImplicitRepresentation, bldg:lod3MultiSurface* down to *bldg:lod4ImplicitRepresentation* and *bldg:lod4MultiSurface*. Note that mappings are possible between the ifcOWL and the cityGML ontologies [205] as well as with the geoSPARQL ontology.

Finally, the Friend-Of-A-Friend ontology or FOAF is a semantic model for describing persons, their activities and their relations to people and objects in the digital world [619]. It can be useful in the USA ontology as to semantically identify the active agents who are in charge of certain aspects. Indeed, some KPIs can be linked to certain interventions that need to be performed in order to improve them and in turn, those interventions can relate to an agent or a service organisation. Therefore the FOAF ontology allows the provision of digital information of the service provider. The core of the ontology contains classes such as *foaf:Organisation*, *foaf:Agent* or *foaf:Person,* their real-world object properties such as *foaf:name, foaf:age, foaf:knows, foaf:member* and their digital world object properties such as *foaf:mbox*, *foaf:workplaceHomepage* or *foaf:currentProject.*

5.3 THE USA ONTOLOGY SCHEMA

As shown in the previous Section 5.2.4, the USA ontology can be divided into sub-modules according to the intended uses and resources reused. Figure 5-4 gives a schematic representation of those modules and how they are structured and linked to one another. The observation module, primarily based on the SSN ontology, is at the centre of the USA ontology. It captures the concepts of sensors and their observations. It is aligned with the KPI module by considering that *usa:Indicator* is a subclass of the *ssn:SensorOutput* and that the *ssn:Observation* must satisfy (*usa:satisfies*) some *usa:SustainabilityGoals.* The KPI module contains the USA structure from themes to indicators as well as their benchmarks. Those are then connected to the QUDT ontology in order to assign them a unit (*qudt:unit*). The Urban Object Module is linked to the Observation Module as *ssn:FeatureOfInterest* is a subclass of *dul:Object*. Additionally, *usa:Intervention* can change (*usa:changes*) the property of certain feature of interest. Finally, entities from the Urban Objects Module and the

Observation Module can be linked to a *ssn:Position* and *ssn:Observation* linked to *ssn:ResultTime,* aligning those modules to the Spatiotemporal Module. Detailed representations of those modules are given in the following sections.

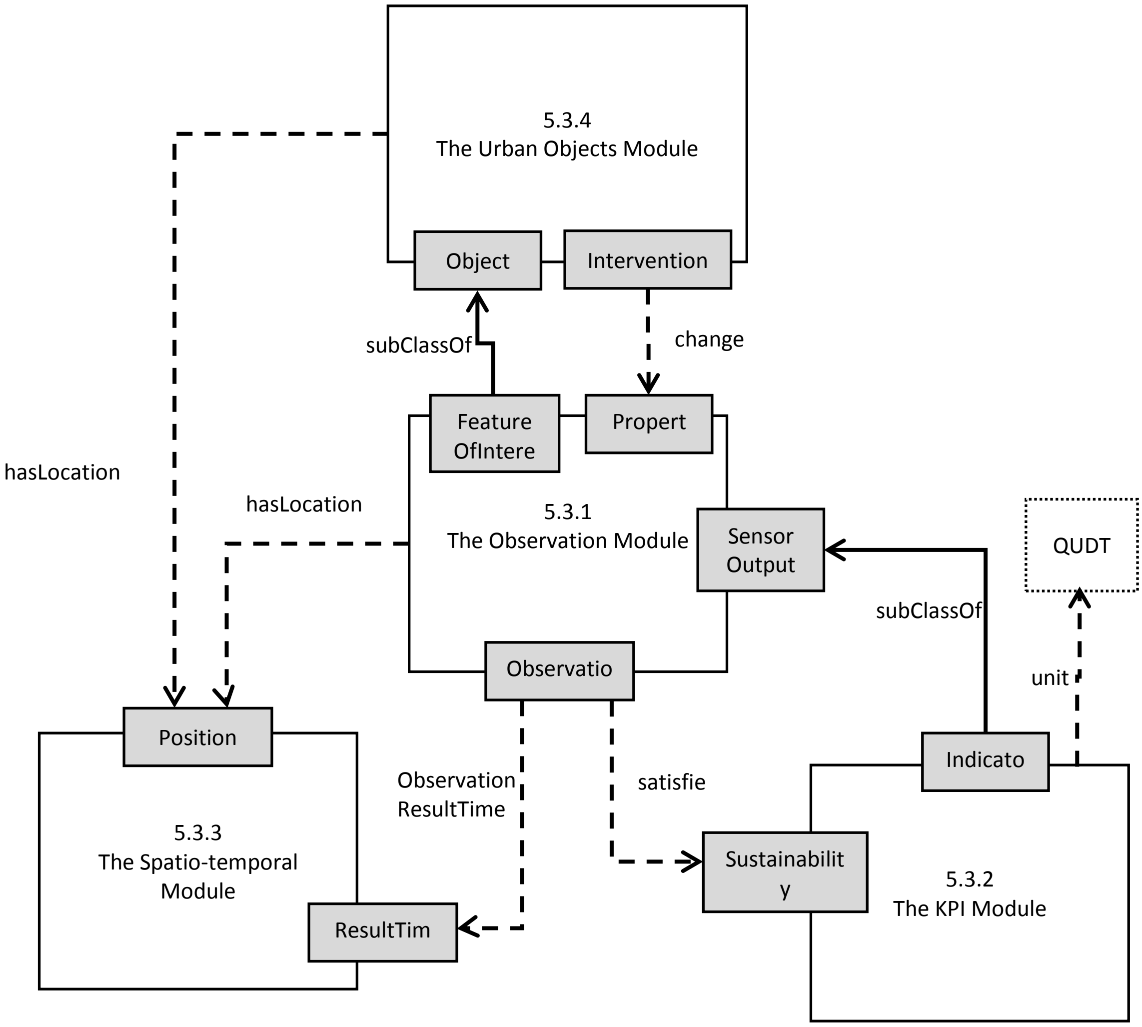


Figure 5-4 USA ontology modular schema

Note that the following sections contain some expressions written using Description Logic syntax. The syntax used is briefly explicated below:

→ ⊑ correspond to the concept of inclusion (is subclass of)

→ ≡ correspond to the concept of equivalence (is equivalent of)

→ ⊓ correspond to the intersection operator (AND)

→ ⊔ correspond to the union operator (OR)

→ ∃ correspond to an existential restriction (SOME exist in)

5.3.1 The Observation Module

Figure 5-5 gives a more detailed description of the observation module. This module is essentially based on the SSN ontology and can be seen as the core of the USA ontology. Indeed, this module allows the semantic modelling of sensors and their observations which ultimately will bridge real-world phenomenon to abstract KPIs. The triple *ssn:Observation*, *ssn:FeatureOfinterest* and *ssn:Property* catch the core concept of the act of observing and the object of this observation. An observation is a situation where the property of a certain feature of interest is observed following a sensing method.

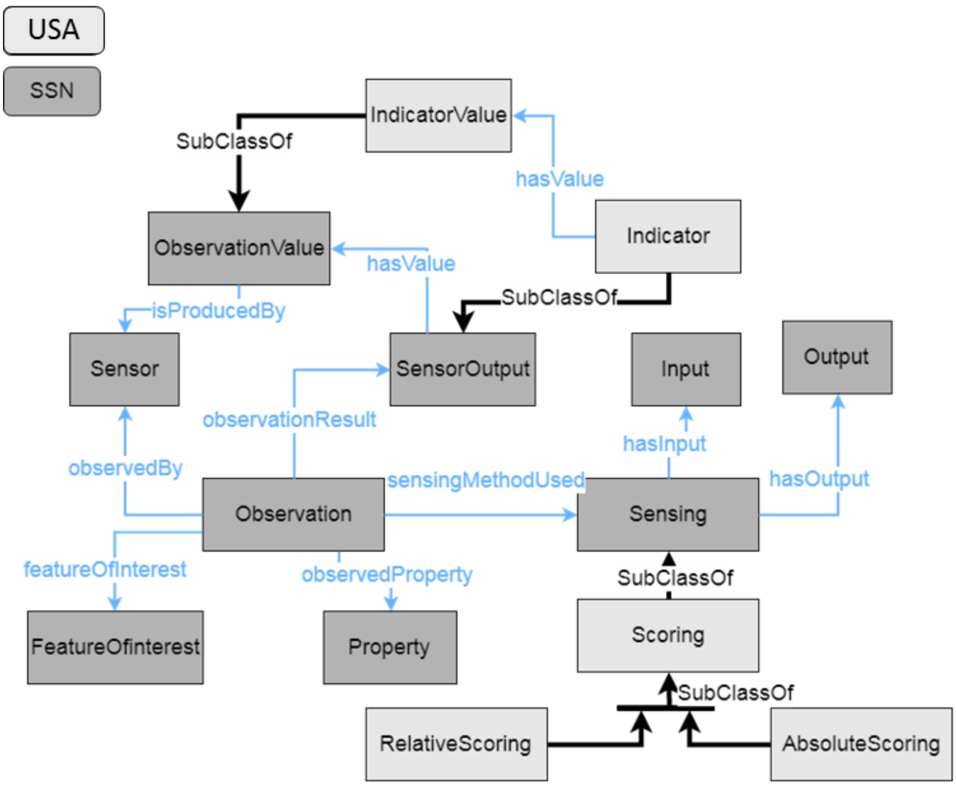


Figure 5-5 USA observation module

In the present case,

> "a *sensor is any entity that can follow a sensing method and thus observe some Property of a FeatureOfInterest. Sensors may be physical devices, computational methods, a laboratory setup with a person following a method, or any other thing that can follow a Sensing Method to observe a Property*." [322]

A sensing method is represented by the class *ssn:Sensing* which is defined as:

> *"a process that results in the estimation, or calculation, of the value of a phenomenon"*

Following this definition, the *ssn:Sensing* allows the inclusion of the *usa:Scoring* method, useful to represent the calculation of the themes, subthemes, criteria and indicators scores. It can be done in an absolute fashion (*usa:AbsoluteScoring*) when compared to benchmarks or relative fashion (*usa:RelativeScoring*) when compared to previous or out-located values. Additionally, the *ssn:Sensing* is linked to its *ssn:Inputs* and *ssn:Output* which allows determining which entities participate in the scorable elements scoring.

In DL terms,

→ *usa:Scoring ⊑ ssn:Sensing;*

→ *(usa:AbsoluteScoring ⊔ usa:RelativeScoring) ≡ usa:Scoring*
where usa:AbsoluteScoring ⊑ (∃ ssn:hasOuput.usa:AbsoluteScore) ;

→ *(usa:SpatialRelativeScoring ⊔ usa:TemporalRelativeScoring) ≡ usa:RelativeScoring*
where usa:RelativeScoring ⊑ (∃ ssn:hasOutput.usa:RelativeScore);

Not represented in Figure 5-5, the USA ontology also contains the following axiom:

→ *(usa:Assessing ⊔ usa:Forecasting) ≡ usa:TemporalRelativeScoring.*

Moreover, the *usa:Indicator* is a subclass of the *ssn:SensorOuput* which is defined as a piece of information. This piece of information is uncoupled with its value that is described by the *ssn:ObservationValue.* In the same way, *usa:Indicator* is therefore uncoupled with its value described by *usa:IndicatorValue*, a subclass of *ssn:ObservationValue*.

In DL terms:

→ *usa:Indicator ⊑ (ssn:SensorOutput ⊓ ∃ ssn:hasValue.usa:IndicatorValue)*
where usa:IndicatorValue ⊑ ssn:observationValue.

Finally, starting from the indicators themselves, different features of interest and properties can be identified and integrated as subclasses. For instances, *usa:TotalEnergyDemandIndicator* results from the observation of the *usa:EnergyProperty* of the *usa:TotalEnergyDemand* feature of interest or *usa:CarbonEmissionsIndicator* results from the observation of the *usa:ConcentrationProperty* of the *usa:CO2Emitted* feature of interest.

In DL terms:

- → *usaTotalEnergyDemandIndicator ≡ (usa:Indicator ⊓ ∃ usa:isObservationResultOf (ssn:Observation ⊓ (∃ ssn:featureOfInterest.usa:TotalEnergyDemand) ⊓ (∃ ssn:observedProperty.usa:EnergyProperty)))*
- → *usaCarbonEmissionsIndicator ≡ (usa:Indicator ⊓ ∃ usa:isObservationResultOf (ssn:Observation ⊓ (∃ ssn:featureOfInterest.usa:Co2Emitted) ⊓ (∃ ssn:observedProperty.usa:ConcentrationProperty)))*

Consequently, efforts have been done to decompose each of the 193 indicators in a feature of interest-property pair. This process resulted in the creation of 187 different objects feature of interest and 37 different property' subclasses.

5.3.2 The KPI Module

Figure 5-6 describes the KPI module including the scorable elements (the 8 themes, 25 subthemes, 90 criteria and 193 indicators). Scorable elements are described in the ontology as:

- → *(usa:Theme ⊔ usa:SubTheme ⊔ usa:Criteria ⊔ usa:Indicator) ≡ usa:ScorableElement.*

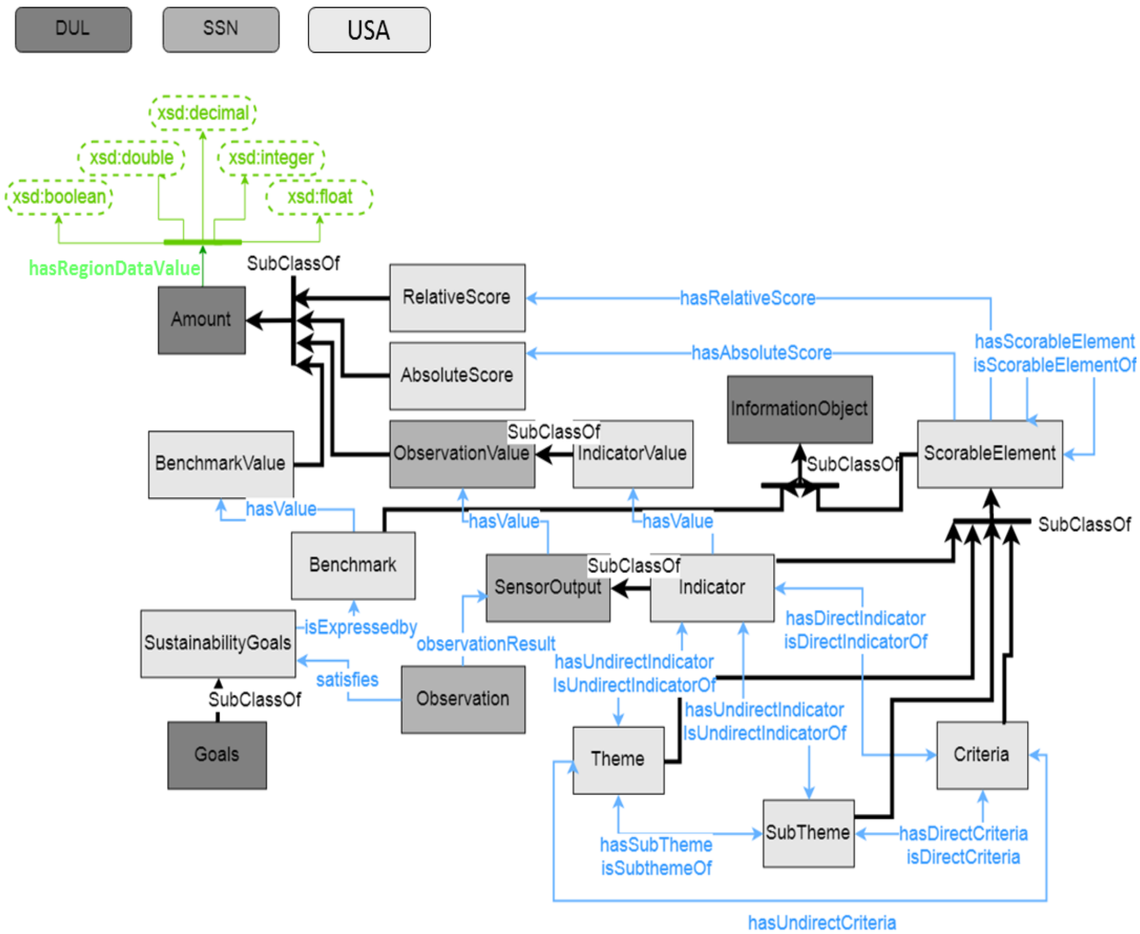


Figure 5-6 USA framework module

The relationships between each element are also introduced. Those relationships must be defined as to answering the competency questions developed in Table 5-1. Consequently, each indicator inclusion into a criterion, and in turn, criteria into subthemes and subthemes into themes must be described. Table 5-4 summarises the existential restrictions that exist between the different scorable elements. In this table, the existential restriction can be described as below:

- → *(usa:isDirectIndicatorOf ⊔ usa:isIndirectIndicatorOf) ≡ usa:isIndicatorOf;*
- → *(usa:hasDirectIndicator ⊔ usa:hasIndirectIndicator) ≡ usa:hasIndicator;*
- → *(usa:isDirectCriteriaOf ⊔ usa:isIndirectCriteriaOf) ≡ usa:isCriteriaOf;*
- → *(usa:hasDirectCriteria ⊔ usa:hasIndirectCriteria) ≡ usa:isCriteriaOf.*

In that manner, one can not only find out which indicator is contained in which criteria but also which indicator is contained in which subthemes or themes. The same goes for criteria and themes relationships and their inverse.

Table 5-4 Existential Restriction between scorable elements

	usa:Indicator	usa:Criteria	usa:SubTheme	usa:Theme	usa:ScorableElement
usa:Indicator	-	usa:isDirect IndicatorOf	usa:isIndirect IndicatorOf	usa:isIndirec t IndicatorOf	usa:isIndicatorOf
usa:Criteria	usa:hasDirect Indicator	-	usa:isDirect CriteriaOf	uusa:isIndire ct CriteriaOf	Usa:isCriteriaOf
usa:SubThem e	usa:hasIndirect Indicator	usa:hasDirect Criteria	-	usa:isSubTh emeOf	usa:isSubThemeOf
usa:Theme	usa:hasIndirect Indicator	usa:hasIndirect Criteria	usa:hasSubThe me	-	-
usa:Scorable Element	usa:hasIndicato r	usa:hasCriteria	usa:hasSubThe me	-	usa:hasScorableElem ent usa:isScorableElemen tOf

Overall, the "simple" description of the scorable elements and their relationships accounts for 321 classes, 8 object properties, 1216 axioms and 890 logical axioms.

Moreover, *usa:Indicator* being a subclass of *ssn:SensorOuput*, itself a subclass of *dul:InformationObject*, all the *usa:ScorableElement* are by extension subclasses of *dul:InformationObject*. Similarly than *usa:Indicator* is linked to *usa:IndicatorValue*, the *usa:ScorableElement* is linked to a type of *usa:Score*, either *usa:RelativeScore* or *usa:AbsoluteScore*, both produces of the *usa:RelativeScoring* and *usa:AbsoluteScoring* described in the previous section.

In DL terms:

→ *ScorableElement ⊑ (dul:InformationObject ⊓ ∃ usa:hasScore.usa:Score)*
where *(usa:hasRelativeScore ⊔ usa:hasAbsoluteScore) ≡ usa:hasScore*
And *(usa:RelativeScore ⊔ usa:AbsoluteScore) ≡ usa:Score.*

Note the *usa:hasScore* and *usa:Score* are not present in Figure 5-6 for better illustrative purposes.

The definition of absolute score requires the inclusion of benchmarks references within the ontology. In DUL, the notion of *Goal* is introduced as

> *"the description of a situation that is desired by an agent, and usually associated to a plan that describes how to actually achieve it"* [324]

This proved to be particularly handy for the description of benchmarks that can be seen as the expression of sustainability goals that the observation must ideally satisfy. The *usa:Benchmark* is *dul:InformationObject* and its value expressed by *usa:BenchmarkValue.*

In DL terms:

→ *usa:SustainabilityGoals ⊑ dul:Gaols;*
→ *usa:SustainabilityGoals ⊑(∃ dul:isExpressedBy.usa:Benchmark);*
→ *ssn:Observation ⊑(∃ dul:satisfies.usa:SustainabilityGoals).*

Finally, all the value related classes namely *usa:IndicatorValue*, *usa:BenchmarkValue*, *usa:RelativeScore* and *usa:AbsoluteScore* are subclasses of the *dul:Amount,* a class used to describe quantities and linked to the actual value (boolean, decimal, float, integer or double) via the data property *dul:hasRegionDataValue*.

5.3.3 The Spatio-temporal Module

Figure 5-7 shows the Spatio-temporal module which helps to give a location to the different entities and a specific time to the observations. An observation can be seen as a single point within a time series. Consequently, each observation is assigned a *usa:ResultTime*, a *usa:SamplingStartTime* and a *usa:SamplingEndTime* via the object properties *ssn:observationResultTime* and *ssn:observationSamplingTime*. SamplingStartTime and SamplingEndTime correspond to the times that frame the validity of an observation (often the time in between two logs) whereas ResultTime is the time that is registered when the observation value is acquired (often the timestamp of the observation). Those classes are considered as subclasses of the *dul:TimeInterval* class used to describe the time dimension via the data property *dul:hasRegionDataValue*. They were not originally part of the SSN ontology and have been added in order to differentiate the different times uses.

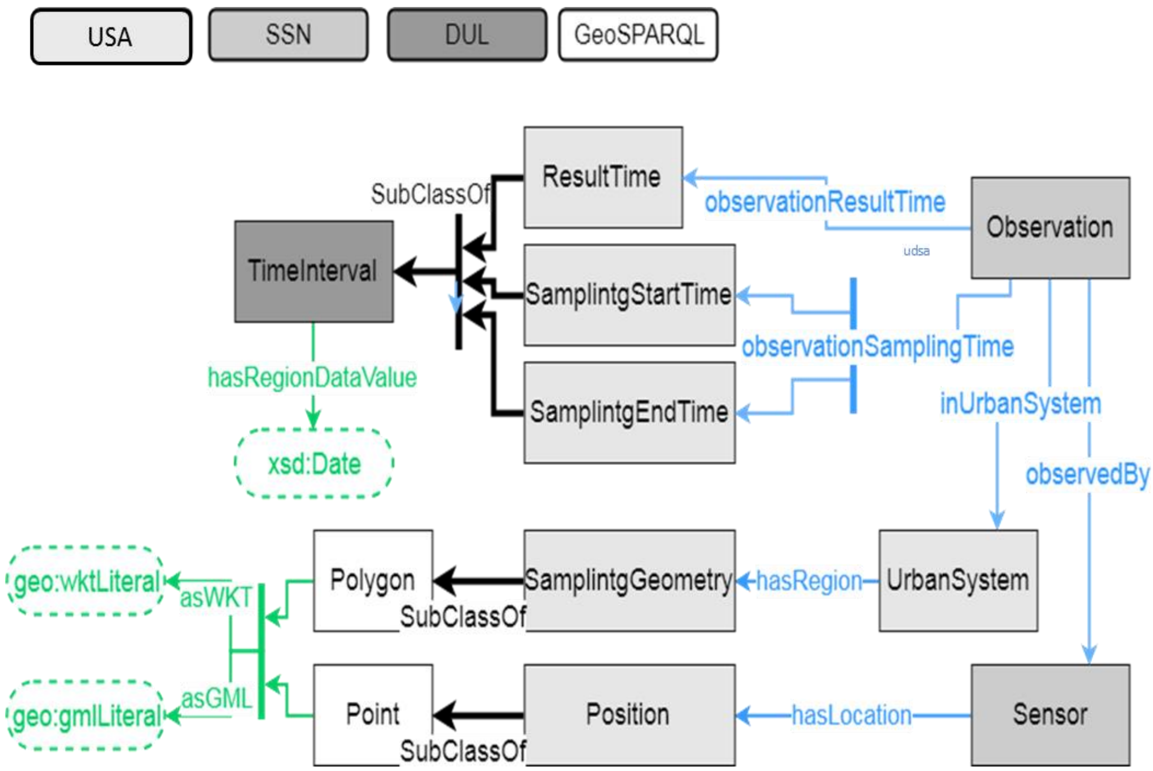


Figure 5-7 USA Spatiotemporal module

In term of space, an observation is considered to be valid within the boundaries of a certain *usa:UrbanSytem*. The object property *usa:inUrbanSytem* expresses this membership. The urban system itself is defined by a *usa:SamplingGeomerty* which is a subclass of *sf:Polygon.* The *sf:Polygon* class is part of the GeopSPARQL ontology and is defined as:

> *"a planar Surface defined by 1 exterior boundary and 0 or more interior boundaries. Each interior boundary defines a hole in the Polygon."* [353]

It can then be expressed in WKT or GML.

In DL term:

→ ssn:Observation *⊑(∃ usa:inUrbanSystem.usa:UrbanSystem);*

→ *usa:UrbanSystem ⊑(∃ dul:hasRegion.usa:SamplingGeometry)*
where *usa:SamplingGeometry ⊑ sf:Polygon;*

→ *sf:Polygon ⊑(∃ geo:asWKT.geo:WKTLiteral)*
or *sf:Polygon ⊑(∃ geo:asGML.geo:GMLLiteral).*

Similarly, a *ssn:Sensor* can be pinned to a particular *usa:Position* via the object property *dul: hasLocation*. The *usa:Position* is a subclass of *sf:Point* which is defined in the GeoSPARQL ontology as:

> *"a Point is a 0-dimensional geometric object and represents a single location in coordinate space".* [353]

It can also be expressed in WKT or GML.

In DL term:

→ ssn:Sensor$\sqsubseteq$*(∃ dul:hasLocation.usa:Position);*
 where *usa:Position* $\sqsubseteq$ *sf:Point;*

→ *sf:Point* $\sqsubseteq$*(∃ geo:asWKT.geo:WKTLiteral)*
 or *sf:Point* $\sqsubseteq$*(∃ geo:asGML.geo:GMLLiteral).*

Note that the SamplingGeometry and Position are here assigned to the UrbanSytem and Sensor classes for illustrative purposes. In reality, no restrictions apply and those could be assigned to any entities that exist in the spatiotemporal dimensions.

5.3.4 The Urban Objects Module

Figure 5-8 presents the Object module where real-world objects are represented with the help of the class *dul:Object*. The class *dul:Object* is defined as:

> "*Any physical, social, or mental object, or a substance. Following DOLCE Full*, objects are always participating in some event (at least their own life), and are spatially located." [325]

Consequently, every entity participating in the assessment can be interpreted as being a subclass of *dul:Object* such as *usa:urbanSytem, ssn:FeatureOfInterest, dul:InformationObject, dul:Agent and dul:Action.* Same goes for *ifc:IfcObject* and *core:CityObject* that can be considered a subclass of *dul:Object*.

Not shown in Figure 5-8 for illustrative purposes but important, some of the 187 object classes derived from the KPIs in Section 5.3.2 can be aligned with the IfcOWL and cityGML ontologies defined object. For instance, *usa:AccessibleCrossWalk* (feature of interest of the *usa:AccessibleCrossWalkIndicator*) can be seen as a subclass of *tran:Track* in cityGML or *usa:EquippedBuildings* (feature of interest of the *usa:EquippedBuildingsIndicator)* can be seen as a subclass of *ifc:Building.* In this way, the USA ontology has multiple classes linking the feature of interest to the cityGML and ifcOWL ontologies. Additionally, efforts have been made to breakdown each of the 187 initial objects identified into elementary ones in order to create classes that could be more easily be connected to those ontologies.

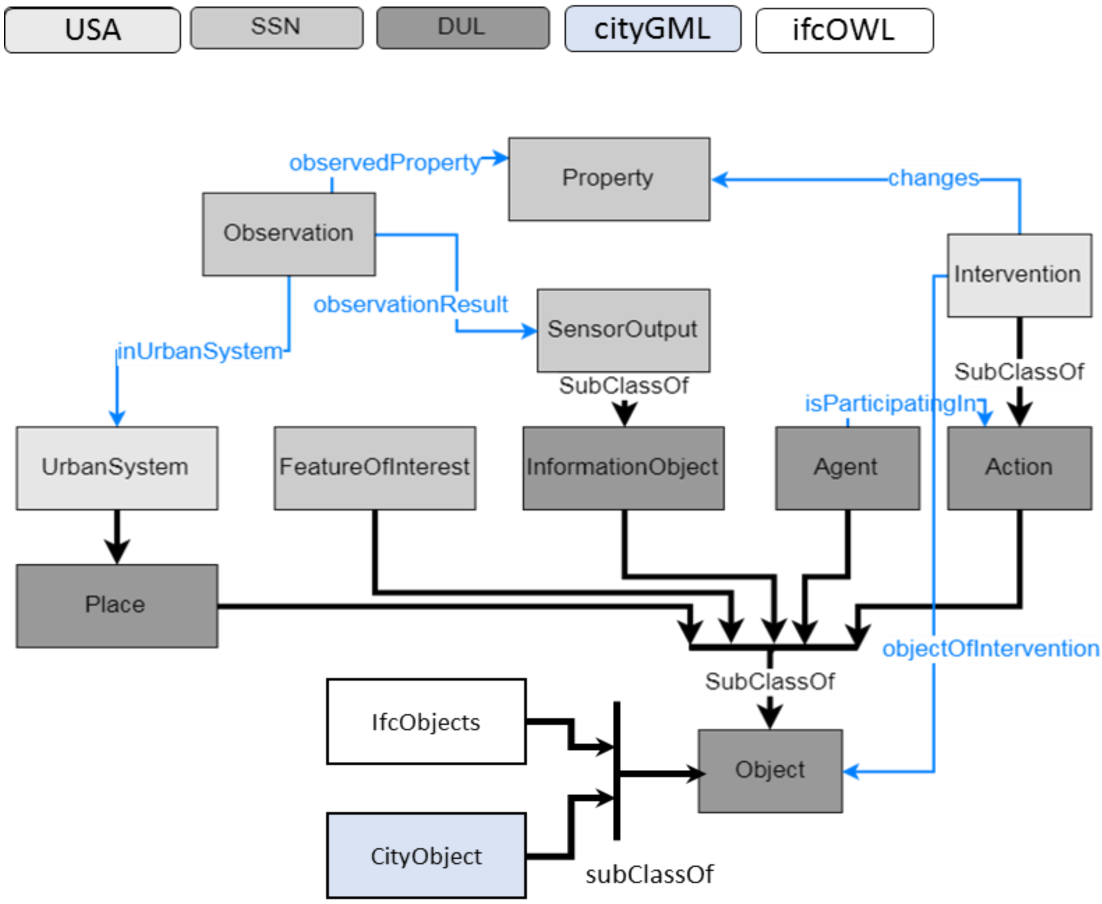


Figure 5-8 USA urban object module

For instance:

→ *usa:EquippedBuilding ⊑ (usa:Building ⊓ ∃ dul:hasComponent. usa:CommunicationTechnology))*

where *usa:Building ≡ ifc:IfcBuilding ≡ bldg.:Building*

and *usa:CommunicationTechnology ≡ (ifc:IfcDistributionControlElement ⊔ ifc:IfcCommunicationsAppliance)*

This approach has led to the creation of around 234 new classes and 14 new object properties.

Furthermore, Figure 5-8 shows the *usa:Intervention* class that is an action (*dul:Action*) done by an agent (dul:Agent) that will modify some properties observed of a feature of interest.

Consequently, one can track changes in properties and relate those to possible actions and people.

5.4 TEST ON THE TBOX

In this section, the ontology is tested against a set of DL queries that showcase how one can explore the USA scheme using the implemented axioms. Figure 5-9 gives an example of DL query in Protégé that have been executed using the ELK 0.4.3 reasoner on a desktop computer with 1TB HDD, Intel Core i7-4790 CPU 3.60GHz, 24 GB memory and Windows 7 64-bits. In this example, all the subclasses of the Indicator class are retrieved. Note that Figure 5-9 has been cropped and not all the subclasses are displayed here.

Those queries follow the competency questions described in Table 5-1 where the links between the different scorable elements present in the USA scheme is queried (e.g. Which scorable element has criteria X?, Which scorable element has indicator X?). Note that not all the competency questions are displayed as example in this section because of their redundancy.

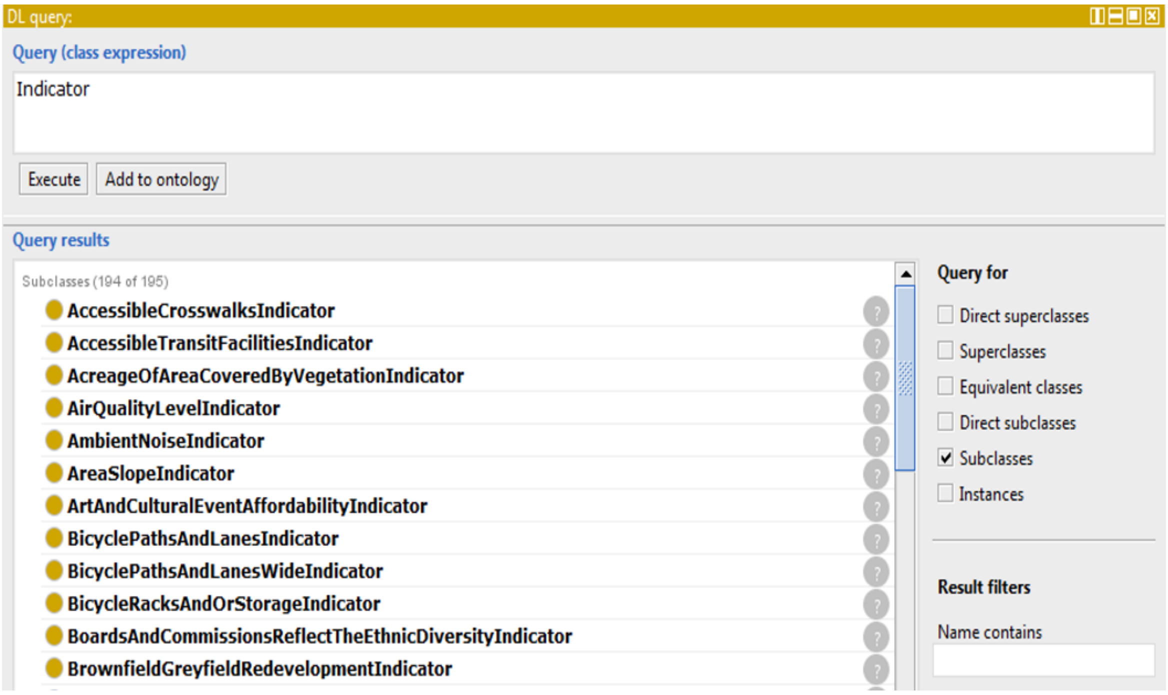


Figure 5-9 Protégé DL query

Query 1. What are the Indicators within the EnergyUseCriteria?

DL query	Output (4 records in 0.114 sec)
`subClassOf isDirectIndicatorOf some EnergyUseCriteria`	ElectricityFromRenewableSourcesIndicator HeatingFromRenewableSourcesIndicator SolarPotentialOnRoofIndicator TotalEnergyDemandIndicator

Query 2. What are the Themes and Criteria that contains the TotalEnergyDemandIndicator?

DL query	Output (3 records in 0.100 sec)
`subClassOf (Theme or Criteria) and hasIndicator some TotalEnergyDemandIndicator`	EnergyUseCriteria InfrastructurePerformanceCriteria ResourcesAndClimateTheme

Additionally, a set of SPARQL queries has been executed in Query 3, in order to test the linkage between the USA native classes and the ifcOWL and cityGML ontologies' classes.

Query 3. What are the relationships between USA entities and ifcOWL/cityGML entities?

SPARQL query

```
SELECT ?subject ?object ?predicate WHERE {
{
	?subject rdfs:subClassOf ?object.
	FILTER regex(str(?subject), "urban").
	FILTER regex(str(?object), "IFC4|opengis").
	BIND("SubClass" AS ?predicate)
}
UNION
{
	?subject rdfs:subClassOf ?object.
	FILTER regex(str(?subject), "IFC4|opengis").
	FILTER regex(str(?object), "urban").
	BIND("SubClass" AS ?predicate)
}
UNION
{
	?subject owl:equivalentClass ?object.
	FILTER regex(str(?subject), "IFC4|opengis").
	FILTER regex(str(?object), "urban").
	BIND("Equivalent" AS ?predicate)
}
}
```

Subject	Predicate	Object
HVAC	SubClass	IFC4_ADD2#IfcDistributionElement
EnergyInfrastructure	SubClass	IFC4_ADD2#IfcDistributionSystem
PublicAppliances	SubClass	citygml/cityfurniture/2.0/CityFurniture
UrbanSystem	SubClass	citygml/cityobjectgroup/2.0/CityObjectGroup
RenewableElectricitySource	SubClass	IFC4_ADD2#IfcEnergyConversionDevice
SamplingGeometry	SubClass	ont/sf#Polygon
UrbanArtPieces	SubClass	citygml/cityfurniture/2.0/CityFurniture
RenewableHeatSource	SubClass	IFC4_ADD2#IfcEnergyConversionDevice
Position	SubClass	ont/sf#Point
GarageDoor	SubClass	IFC4_ADD2#IfcDoor
SolarPanel	SubClass	IFC4_ADD2#IfcSolarDevice
IFC4_ADD2#IfcWasteTerminal	SubClass	WasteCollectionPoint
citygml/building/2.0/Room	SubClass	InternalBuildingsSpace
IFC4_ADD2#IfcRoof	Equivalent	Rooftop
IFC4_ADD2#IfcSpace	Equivalent	InternalBuildingsSpace
citygml/landuse/2.0/LandUse	Equivalent	ActivitySegmentedSpace
IFC4_ADD2#IfcBuilding	Equivalent	Building
citygml/building/2.0/RoofSurface	Equivalent	Rooftop
IFC4_ADD2#IfcWall	Equivalent	Wall
citygml/transportation/2.0/Road	Equivalent	Roadway
citygml/building/2.0/Building	Equivalent	Building
IFC4_ADD2#IfcBuildingElement	Equivalent	BuildingComponent
citygml/vegetation/2.0/_VegetationObject	Equivalent	Vegetation
IFC4_ADD2#IfcConstructionMaterialResource	Equivalent	Materials
citygml/building/2.0/WallSurface	Equivalent	Wall

5.5 SUMMARY

As described in Section 2.3, the engineering of a system that leverage the IoT and other various sources will require the implementation of a technological approach that can deal with their heterogeneity. Indeed, interoperability is a key aspect to achieve holistic and ubiquitous services upon the smart city. Consequently, in the perspective of developing a real-time sustainability assessment scheme that fully takes advantage of sensor networks, one must consider solutions to deal with this challenge. Web semantic technologies benefit from a long time development and a strong community that is demonstrating in practice how they contribute to the creation of interoperable systems. Indeed, ontologies are semantic models of real-world concepts that allow machines to capture information meaning beyond its simple syntax. In this present context, the USA ontology has been developed. The ontology represents the complete USA scheme including all the themes, subthemes, criteria and indicators as well as the entire process for their determination, from the indicator's object to the mean of measurement and the calculation methods themselves. The aim is that any relevant source of information can be introduced in the ontology and linked to the built environment and KPIs.

Additionally, the concept of sustainability itself is complex by nature as it covers several often interconnected domains. The creation of an ontology can help in the understanding of sustainability as it describes accurately the connections between indicators and their environment. One can therefore query the semantic model to better understand which particular aspects are involved in the creation of sustainable places. This comes also as an answer to one of the concerns given in Section 2.1 which is the issue of transparency in current urban sustainability assessment schemes. The USA ontology makes the scheme structure transparent and enables the user to query it.

The development of the USA ontology required a rigorous methodology. This is due to the complex nature of the web semantic modelling itself. Indeed, the main motivation for the development of ontologies is the creation of linked data and unified information space. Therefore, when creating a semantic model of a certain domain, one must consider how it will fit within the already constructed ontological structures. In this prospect, the developer must scope the boundaries of the domain to be modelled and choose the most relevant existing ontological resources to support it. The NeOn methodology details those principles and gives insights on how to proceed for an efficient ontological development. It has been followed for the engineering of the USA ontology. The ontology requirement specification with the help of the competency questions has brought focus on essential concepts that must be included and has drawn terminologies that can help in the ontological resources discovery. The author has identified four main aspects that must be included: the USA scheme itself with its different themes, subthemes, criteria and indicators; the representation of sensors and means to measure the USA KPIs; the possibility to locate an entity and to set a time of validity; and the representation of the object participating in the city dynamic from the built environment to people or organisation.

If the capture of the sustainability indicators is directly taken from the USA scheme created, the other aspects have been subject to ontologies' reuses. The smart city sensors and their readings are capture in the SSN ontology that describes sensor networks, observation, sensing method etc. Geospatial information are modelled within the geoSPARQL ontology. This ontology allows the description of geometries and coordinates and how they relate one to another, and is aligned with known standards such as GML or WKT. Time information are already covered in the SSN and DUL ontologies, describing observations sampling times and result time. The ifcOWL ontology describes buildings, their components and appliances as well as processes and people that engage in their construction and maintenance. Additionally, the cityGML ontology covers the extended built environment with the inclusion of the city's uses, furniture, structures etc.

The USA ontology is then structured around those four aspects in a modular fashion. The Observation module will catch sensors, observation and sensing processes based on the SSN ontology. This module is linked to the KPIs module via the integration of the scorable elements as the outputs of the sensor observations. Those scorable elements are fully described, especially each elements membership into another. The Object module interprets the features of interest of a sensor observation as real-world objects described in the ifcOWL and cityGML ontologies. Finally, the spatio-temporal module allows setting a time to the observation and their results and geospatial references to the different objects.

Overall, the chapter has provided evidences to answer the following research questions:

RQ2. How an effective urban sustainability assessment can help different parties of a city in their decision making?

Semantic web technologies presented in this chapter will influence decision making by introducing meaning to data, understandable by humans and machines. The logical axioms developed in such information system allow one to infer extra piece of knowledge, introducing intelligence into the system. As such, one can more efficiently query and discover information on sustainability and the urban environment which in fine will help decision making.

RQ3. How can sustainability assessment leverage the smart city paradigm, specifically ICTs and the IoT?

The USA ontology has been built around the SSN ontology that semantically describes sensor networks and their observations. Therefore, data produced by divers' sensors can be logically linked to sustainability KPIs. Semantic web technologies will help in the creation of a unify system of information where sustainability KPI will fully leverage on data generated by the smart city and its sensor.

RQ4. How can semantic web technologies unify heterogeneous data resources for holistic services and applications?

By linking several domain ontologies within the USA ontology, semantic web technologies have demonstrated how heterogeneous information could be link despite of their respective format. For instance, a BIM model can be semantically link to sensor data or cityGML model.

The next chapter will focus on the testing of the USA framework with the design and implementation of a web service for real-time sustainability assessment. Implemented over actual pilot sites, the service will leverage on the ontology created for the query and retrieval of the KPIs and insightful information related.

The Framework Testing: Web Service for Real-time Sustainability Assessment

6.1 OVERALL DESCRIPTION

Chapter 4 and 5 have introduced the basis of the USA framework from the scheme development to the creation of the USA ontology. 8 themes, 25 subthemes, 90 criteria and 193 indicators have been selected based on 29 existing urban sustainability assessment schemes. Efforts have been made in order to semantically engineer an USA ontology that is based on the USA scheme while offering the possibility to perform a large set of essential queries for KPIs discovery and analysis.

In this Chapter, the full designed framework will be detailed including the mechanisms applied to leverage the USA ontology. Then, this framework will be tested against real case studies in order to validate certain functionalities and to demonstrate its use in real-time.

The framework design will follow the three cycles view of design science research presented in Section 3.4.2 where Information Systems are designed in an iterative manner upon an application context that provides concrete requirements and testing. In that way, the design can be evaluated against real cases which allow the growth of expertise and meta-design understanding. In the present case, early development has unveiled technical challenges that required substantial efforts in order that a viable framework comes out. Indeed, there have been two main challenges in the development of the USA framework: the first one relating to scaling the intensive data flow for an efficient system and the second one relating to the semantic alignment between information retrieve from the sensor and the USA ontology. Those will be discussed in the following sections.

After a number of development iterations, the proof of concept in Figure 6-1 has been drawn. The concept can be divided into four main modules: data acquisition and sensors instances storage; the semantic model storage within a triple store; the ONTOP module that bridges the two first modules; and finally the front-end module with dashboards and 3D models. The originality of such system is the division between the “finite” data that are stored in the triple store and the “infinite” data stored in a relational database. Such system is called an Ontology-Based Data Access (OBDA) where a user can access information through ontological queries which are then translated, on-the-fly, into SQL queries to retrieve the corresponding information stored in an relational database [331]. The objectives are threefold [620]:

– *Variety* with the integration of a common schema that unified heterogeneous local data schema
– *Velocity* with on-the-fly data access that allows the capture of fresh data
– *Volume* with large dataset being virtually integrated to the ontology via the mapping with a relational database.

Various studies have already considered the use of OBDA to enhance Information System [331], [620]–[624]. Reasons for using an OBDA will be further described in the following section.

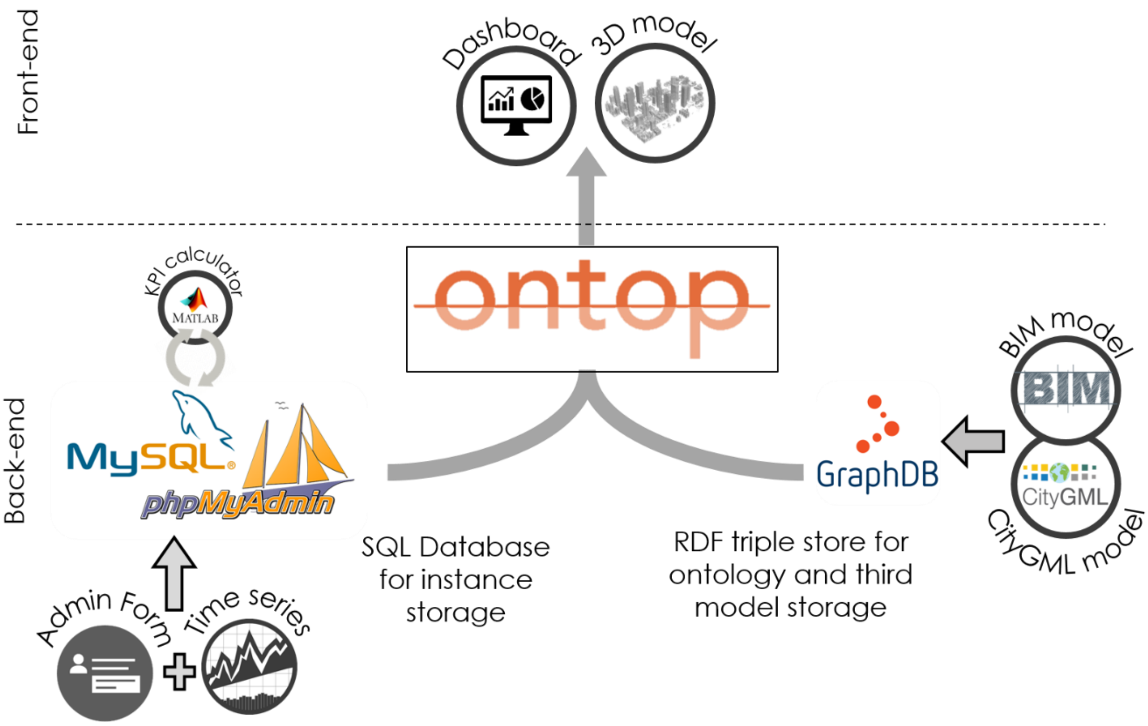


Figure 6-1 USA framework proof of concept

The following research questions will be addressed in this chapter:

RQ2. How an effective urban sustainability assessment can help different parties of a city in their decision making?

RQ3. How can sustainability assessment leverage the smart city paradigm, specifically ICTs and the IoT?

RQ4. How can semantic web technologies unify heterogeneous data resources for holistic services and applications?

RQ5. How technological, human and financial assets relate to such service provision approach in the smart city paradigm?

6.2 WHY AN OBDA?

The reason for the selection of an OBDA has to do with the nature of the web semantic modelling itself and how the information is treated. Indeed, web semantic modelling is an intelligence information system ruled by the implementation of logical axioms as mentioned in Section 2.3.2.

Those axioms not only allow the development of a conceptual framework but also connect it to real-world instances. In web semantic, the knowledge-based is divided into two different components: the TBox and the Abox (Figure 6-2).

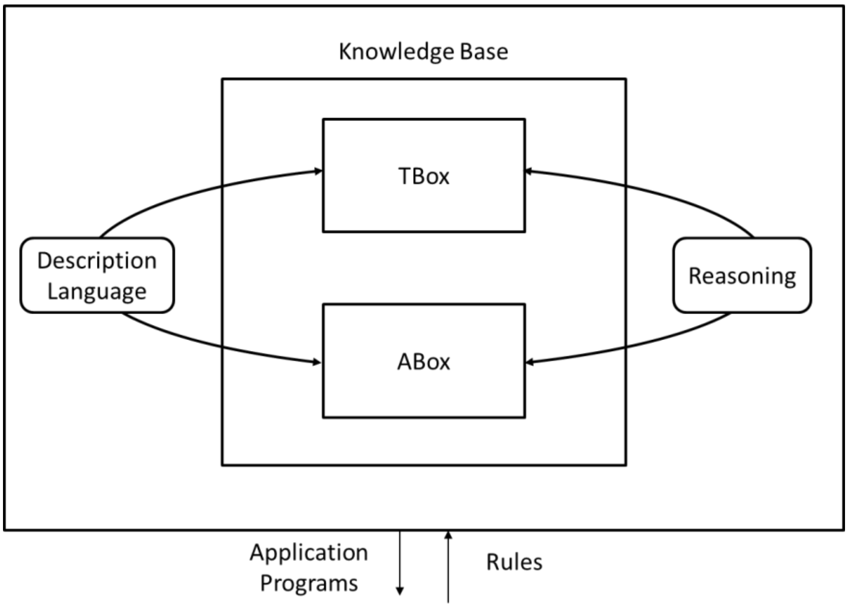


Figure 6-2 The Architecture of Knowledge Representation System [304]

The TBox includes the terminology and the domain description with different concepts and roles while the ABox consists of the individuals' assertions in term of this vocabulary [304]. Table 6-1 gives a concrete example of the difference between the TBox and ABox.

Table 6-1 TBox and ABox example

TBox	ABox
Female ⊑⊤⊓¬ Male	Human(Meryl)
Male ⊑⊤⊓¬ Female	Female(Meryl)
Animal ≡ Male ⊔ Female	Woman(Audrey)
Human ⊑ Animal	Human(Robert)

Woman ≡ Human ⊓ Female	¬ Female(Robert)
Man ≡ Human ⊓¬ Female	Man(Dustin)
Mother ≡ Woman ⊓ ∃hasChild	hasChild(Audrey, Meryl)
Father ≡ Man ⊓ ∃hasChild	hasChild(Robert, Dustin)

The TBox and Abox are given semantics with the definition of interpretations. An interpretation *I* is composed of a non-empty set of individuals $\triangle^{\mathcal{I}}$ (also called domain) and an interpretation function $\cdot^{\mathcal{I}}$ that maps individuals, concepts and roles to the domain. On The TBox, the interpretation *I* satisfies an axiom $C \sqsubseteq D$ or $C \equiv D$ if and only if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ and $C^{\mathcal{I}} = D^{\mathcal{I}}$ respectively. If and only if the interpretation *I* satisfies all the axioms of the TBox, then it is a model of the TBox. On the ABox, the interpretation *I* satisfies a concept assertions $C(a)$ and a role assertion $R(a, b)$ if and only if $a^{\mathcal{I}} \in C^{\mathcal{I}}$ and $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in R^{\mathcal{I}}$. If and only if the interpretation *I* satisfies all assertion of the ABox, then it is a model of the ABox. Finally, a model of both the TBox and ABox is known as an abstraction of a concrete world [304], [625].

The interpretation and model are important for the reasoning process that aims at inferring hidden knowledge. Indeed, when running, the reasoner will proceed to a series of checks on satisfiability, subsumption, equivalence, disjointness and the consistency of the ABox w.r.t the TBox. The interpretation and model theory will determine the valid conclusion or entailment which influences the reasoning complexity.

In semantic web modelling, several entailment regimes exist that interpret a certain vocabulary. Already introduced in Section 2.3.2.2 the current entailment regimes are: RDF entailment, RDFS entailment, D-Entailment, OWL 2 RDF-Based Semantics entailment, OWL 2 Direct Semantics entailment, and RIF-Simple entailment. For instance, it is stated that an RDF graph G RDFS-entails an RDF graph H if every RDFS-interpretation which satisfies every triple of G also satisfies every triple of H. However, the converse is not necessarily true.

To summarise, for a given graph with a given vocabulary, several interpretations are possible which will result in different entailments. Following the complexity of the interpretation, the reasoning process will be influenced leading to more or less decidable solutions.

Decidability is, therefore, a challenge when considering computing all interesting logical conclusions. Because computational power is limited in practice, semantic models must be tractable. Decidability can be measured and is described in classes. From the most complex one down to the simplest ones [306]:

- N2EXPTIME problems are solvable in time that is at most double exponential in the input size;
- NEXPTIME problems are solvable in time that is at most exponential in the input size;
- PSPACE problems are solvable using space that is at most polynomial in the input size;
- NP problems are solvable by a nondeterministic algorithm in time that is at most polynomial in the input size ;
- PTIME problems are solvable by a deterministic algorithm in time that is at most polynomial in the input size;
- LOGSPACE problems solvable using space that is at most logarithmic in the input size.

Table 6-2 gives the complexity level of each OWL 2 and OWL entailment regimes. The OWL languages (OWL 2 RDF-Based Semantics, OWL 2 Direct Semantics and OWL 1 DL) are unfortunately in the worst case highly intractable. Indeed, problems with a complexity above PTIME are often referred as intractable. Therefore, a viable solution has been to limit the expressivity of the OWL 2 language by introducing tractable fragments such as the OWL 2 EL, OWL 2 QL and OWL 2 RL [306]. Shown in Table 6-2 , those fragments are LOGSPACE to PTIME-complete.

Table 6-2 Entailment complexity [306]

Language	Reasoning Problems	Combined Complexity
OWL 2 RDF-Based Semantics	Ontology Consistency, Class Expression Satisfiability, Class Expression Subsumption, Instance Checking, Conjunctive Query Answering	Undecidable
OWL 2 Direct Semantics	Ontology Consistency, Class Expression Satisfiability, Class Expression Subsumption, Instance Checking	N2EXPTIME-complete
	Conjunctive Query Answering	Decidability open
OWL 2 EL	Ontology Consistency, Class Expression Satisfiability, Class Expression Subsumption, Instance Checking	PTIME-complete
	Conjunctive Query Answering	EXPTIME
OWL 2 QL	Ontology Consistency, Class Expression Satisfiability, Class Expression Subsumption, Instance Checking	NLogSpace-complete
	Conjunctive Query Answering	NP-complete

	Ontology Consistency	PTIME-complete
OWL 2 RL	Class Expression Satisfiability, Class Expression Subsumption, Instance Checking	co-NP-complete (PTIME-complete for atomic class expressions)
	Conjunctive Query Answering	NP-complete
OWL 1 DL	Ontology Consistency, Class Expression Satisfiability, Class Expression Subsumption, Instance Checking	NEXPTIME-complete
	Conjunctive Query Answering	Decidability open

Back to the USA framework, the USA ontology is constructed based on the SSN ontology where each and every point in a time series is considered as an *ssn:Observation*. This means that there will be at least as many triples as there are time points in a time series. Consequently, time series being by nature "infinite", the investigated semantic model will be based on a highly intensive set of individuals and assertions which could lead to undecidable reasoning tasks over the knowledge base. The most valuable solution here is, therefore, the use of the OWL 2 QL fragment that has been designed in order to deal with very large volumes of instances. The model will result in a limited expressiveness but the most important reasoning tasks will be performed in LOGSPACE with respect to the size of the data (assertions). An OBDA relies on the OWL 2 QL entailment and allows the reasoning on the ABox to be scaled down by producing on-the-fly SQL queries over a relational database. Consequently, it complies with the data-intensive nature of the intended application enabling reasoning tasks in decidable computing time. This feature has motivated it use in this research.

6.3 RELATIONAL DATABASE, SENSOR DATA ACQUISITION AND KPI CALCULATION

6.3.1 Relational Database

Relational database management systems (RDBMS) are often seen as the "traditional" type of databases [626]. They are currently the predominant systems to store data for business applications [627]. A relational database is organised as a collection of relations between entities that can be organised in a set of tables with rows and columns [628]. Each cell contains a single value while columns' names are attributes. Attributes and table names must be uniquely identified as to create unambiguous tuples. Note that an attribute can be present in different tables but must be unique within a single table. Tables are connected to one another via primary and foreign keys that map attributes across tables. RDBMSs support the ACID (Atomicity, Consistency, Isolation and Durability) properties which aim to ensure that all the transactions are reliably processed [629]. The programming language behind most RDBMS is a form of Structured Query Language (SQL) which allows the construct, update of the database as well as the retrieval of information. Well-known RDBMS that use SQL are DB2, Microsoft SQL Server, PostgreSQL, MySQL, Oracle etc [630]–[633].

For this research, the MySQL RDBMS has been chosen. This choice is motivated by:

- its gratuity;
- its large community, intensive resources and in fine strong technical support;
- the provision of phpMyAdmin, a free web tool to efficiently manage MySQL databases;
- its easy installation via an installer provided free-of-charge by Bitnami.

As part of the implementation of the USA framework, the MySQL database has been structured based on the database used in the 52°North project [634], a project that focuses on the development of collaborative geospatial open source software. More specifically, the database schema is partially based on the 52°North Sensor Observation Service (SOS) application, an "*interoperable web-based interface for inserting and querying sensor data and sensor descriptions*" [634]. The service is an implementation of the research undertaken by the OGC and described in the OGC® Sensor Observation Service Interface Standard [635]. The service is particularly interesting in the present research as it follows the OGC standards [319], [320], [635]–[638] that also feature the SSN ontology, used in this research [323]. Therefore the 52°North database schema is aligned with the SSN ontology schema. The entire database model of the 52°North Sensor Observation Service (SOS) can be found in the online wiki documentation [639]. The iterative methodology has led to some modification of the database schema in order to accommodate certain features needed.

Figure 6-3 shows the database schema of USADB, the MySQL database that runs within the USA framework. For convenience, the depicted schema is a trimmed version of the actual schema where the following have been removed:

- the *usa.unit* table that links the *unitId* asserted in other tables with a string *unit* containing the actual unit of measurement;

- the *usa.blobvalue, usa.booleanvalue, usa.countvalue, usa.geometryvalue* and *usa.textvalue* tables that, in a similar fashion to *usa.numericvalue,* link an *usa.observation* to its actual value in the format referred in the table;
- in the same way, *usa.sensorfieldgeometryvalue, usa.sensorfieldtextvalue, usa.sensorcapabilitiesnumericvalue, usa.sensorinputcountvalue* etc*;*
- *usa.sensorkeyword,* a table that allows setting keywords.

The schema can be divided into three portions that cover the admin details, sensor metadata and sensor data. The admin details are centred on the *contact* table while the sensor metadata and sensor data around the s_procedure (aka sensor) and series tables respectively.

Table 6-3 describes in more details the tables' content and their links.

Table 6-3 USADB Tables Description

USADB tables	Description	Links
contact	Contains the contact details of the admin e.g. phone, email, organisation, role etc	role.roleid organisation.organisationid
featureofinterest	Contains the foi details with its identifier, name, open description, position etc	series.featureofinterestid
[X]value	Contains Numeric/Boolean/Geometry/Count … observation value	observation.observationid
observableproperty	Contains the observed property details with its identifier, name and open description.	series.observablepropertiesid sensorinput.inputobservedpropertiesid
observation	Table containing the observation details with its identifier, name, result and validation times, unit, position etc.	series.seriesid unit.unitid [x]value.observationid
organisation	Contains the name and address of the organisation administrating the observation	contact.organisationid
role	Contains the role of the individual administrating the observation	contact.roleid
sensorcapabilities	Contains sensor capabilities details e.g. battery life duration, recorded position list in the case of moving sensor....	s_procedure.procesureid sensorcapabilities[x]value. capabilitiesvalueid unit.unitid
Sensorcapabilities[X]value	Contains Numeric/Boolean/Geometry/Count … value of a sensor capability	sensorcapabilities.capabilitiesvalueid
sensorcontact	Many-to-Many table linking contacts to sensors	contact.contactid s_procedure.procedureid
sensorfield	Contains sensor capabilities details e.g. colour, weight, voltage, amperage ...	s_procedure.procesureid sensorfield[x]value. fieldvalueid unit.unitid
Sensorfield[X]value	Contains Numeric/Boolean/Geometry/Count … value of a sensor field	sensorfield.fieldvalueid

sensorinput	Contains sensor input details such as input label and pointing to the input values used for the observation determination e.g. numeric, text, geometry or simply an observed property	s_procedure.procesureid sensorinput[x]value. inputvalueid observableproperty.observablepropertyid unit.unitid
sensoroutput	Contains sensor input details such as output label e.g. Energy Demand Indicator, Water Consumption Indicator..., unit and description	s_procedure.procesureid unit.unitid
sensorposition	Contains coordinate of the sensor position.	s_procedure.procesureid
sensorsystem	Relation table to store sensors hierarchies.	s_procedure.procesureid
series	Table to store a (time-) series which consists of featureOfInterest, observableProperty, and sensor.	s_procedure.procesureid featureofinterest.featureofinterestid observableproperty.observablepropertyid
s_procedure	Contains sensors details such as identifier, name and open description.	sensorcapabilities.procedureid sensorcontact.procedureid sensorfield.procedureid sensorinput.procedureid sensoroutput.procedureid sensorposition.procedureid sensorsystem.procedureid series.procedureid

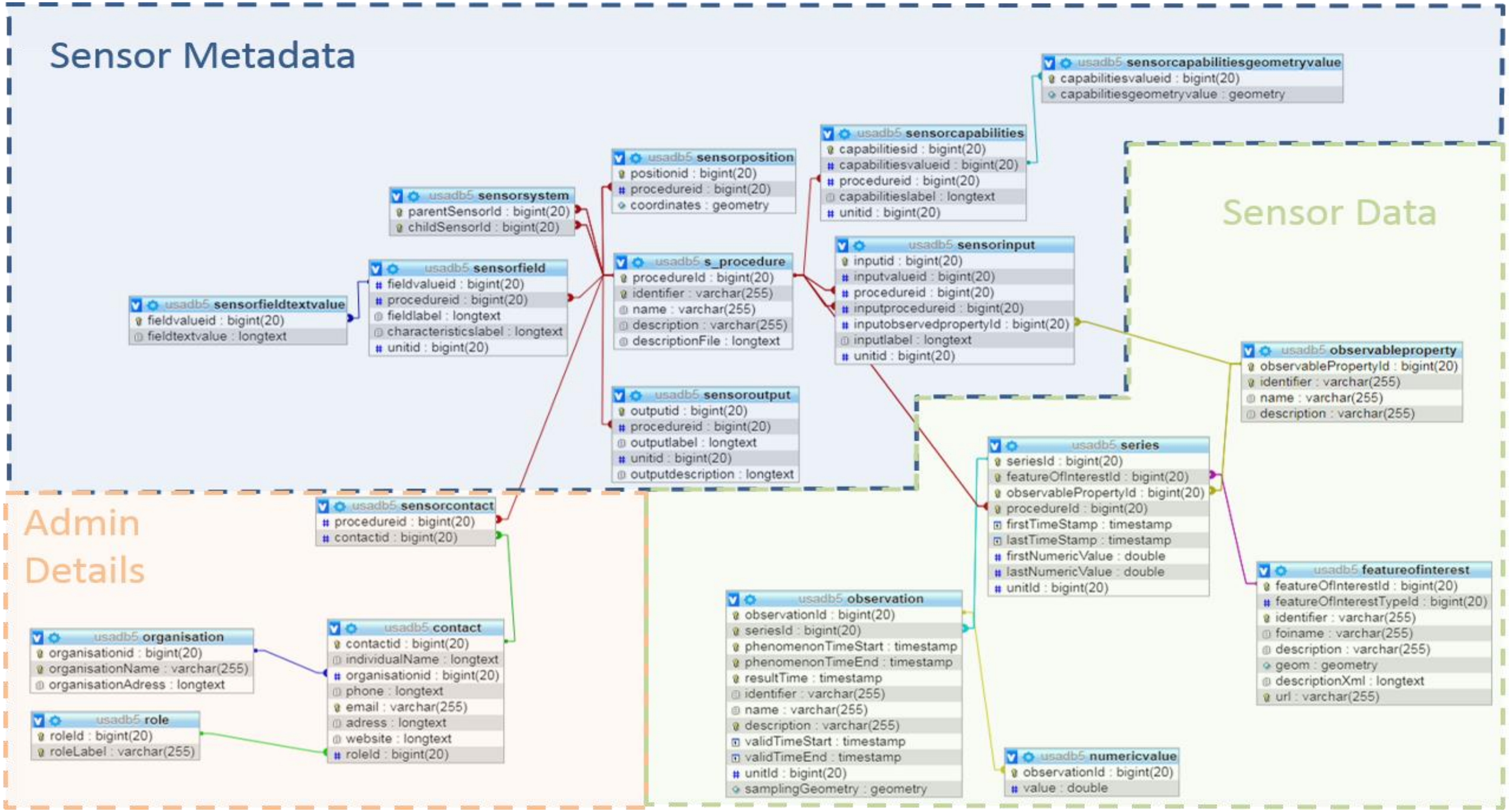


Figure 6-3 USA Database schema sample

6.3.2 Sensor Data Acquisition

The data have been taken from several sensors deployed within the different pilot sites considered for this study already described in Section 3.4.7.

For the site of The Works in Ebbw Vale, a set of sensors has been installed in the different buildings as part of the BEMS system that aims to control district heating and cooling inputs. As shown in Table 4-5 the following information were recorded:

- Outdoor Temperature
- Indoor Temperature
- Buildings Electricity Demand
- Buildings Heat Demand
- Local Electricity Produced
- Local Heat Produced
- Biomass Waste Volume Consumed

In the present case, direct access to the sensor live feed appeared to be impossible due to technical limitation and data protection. Indeed, in Ebbw Vale, the BEMS was provided by a contractor that let the district heating manager access information via a proprietary software. An automated connection to this software by a third party (Cardiff University) was not authorised. Consequently, historical data have been collected manually using this software's UI with the objective to demonstrate how the USA framework could run KPIs calculations and queries. Note that the limited amount of historical data on indoor temperatures did not allow any exploitable development and that they have therefore been simulated. Those historical data stored in a collection of CSV files have been populated automatically into the MySQL database via the Hibernate library in JAVA, an object-relational mapping framework for the Java programming language [640]. The program parses specific information types to send them to the appropriate tables and attributes, creating relationships while ensuring no clashes.

Because of the technical limitations of the Ebbw Vale case study, two other cases have been investigated in order to validate the feasibility of the tool over live data; "52 The Parade" and THERMOSS project pilot sites. The information was then fetched from the APIs using a job builder and scheduler implemented in JAVA that will order a periodic data retrieval. Each new set of information is parsed with the Hibernate library in JAVA and stored in the MySQL database.

Table 6-4 gives an example of the parsing procedure using Hibernate between JAVA objects and MySQL DB attribute. In blue are the required details when retrieving historical data from tabular files while in orange is the parsing details for the retrieval of live data.

Note that these configurations have been implemented in order to have control over the database to make tests. Ideally, an OBDA can be set over any type of relational database and there is no need to parse the information from a database schema to another. Consequently, heterogeneous sources with different schemas can be unified through a single system.

Table 6-4 JAVA objects to SQL DB equivalences

Generic Name	Java Object	USADB Attribute	Description	Example
Role	role_srting	role.rolelabel	Role, Statue of the admin	PhD Student
Organisation	org_string	organisation.organisationname	Admin's organisation	BRE
Email	email_string	contact.email	Admin's email and unique identifier	J.Doe@cardiff.ac.uk
User Full Name	username_string	contact.name	Admin's full name	John Doe
User Address	useradress_string	contact.address	Admin's or organisation's address	52 the parade
User Phone	userphone_string	contact.phone	Admin's phone number	654651654
Website	website_string	contact.website	Admin's or organisation website	BRE.com
Sensor	procedure_string	s_procedure.identifier	Unique sensor identifier	Sensor13
Feature of Interest	foiname_string	featureofinterest.foiname	Feature of Interest name	Electricity_EnergyDemand
Observed Property	obspr_string	Observableproperty.name	Observed property name	EnergyProperty
Feature of Interest position	foigeom_string	featureofinterest.geom	Feature of Interest position coded as WKT object (coordinates)	POINT(-3.2039952278137207 51.776660268454144)
Sensor Position	senspos_string	Sensorposition.coordinate	Sensor position coded as WKT object (coordinates)	POINT(-3.2039952278137207 51.776660268454144)
Urban System Geometry	sensgeom_string	*Sensorcapabilitiesgeometryvalue.* capabilitiesgeometryvalue	Urban system geometry coded as WKT object (polygon of coordinates)	POLYGON((-3.2001328468322754 51.77338480231185,-3.2047462463378906.....))
Unit of Measure	unit_string	Unit.unit	Unit of measure of a time serie	kWh
Feature of Interest	foi_string	Featureofinterest.identifier	Unique Feature of Interest identifier	EbbwVale/LearningZone_Electricity_EnergyDemand
Dataset File address	selectedfile	Observation.resultTime Numericvalue.value	Link to the spreadsheet with all the value of the time series. It is then stored in both observation and numericvalue tables	EbbwVale_LearningZone_Electricity_EnergyDemand.xlsx
Host	host		IP address of the original meters server	102.188.1.0
Port	port		Communication endpoint to the server	3362
Service username	username		Desired username to allow server connection	username

Service Password	pwd		Desired password to allow server connection	password
API GET path	path	Observation.resultTime Numericvalue.value	Path can be built from different attribute such as sensorID, timeStart, timeEnd, timeInterval … Depend on the API specifications Value are then stored in the database	http://[IPAddress]/TrendSamples?trendId=03%2FServer%.....

6.3.3 KPI Calculator

6.3.3.1 *Calculation Procedure*

Once the sensors data and meta-data are stored into the RDBMS, the USA framework will calculate the appropriate KPIs corresponding to those. Figure 6-4 gives an overview of the KPIs calculation procedure.

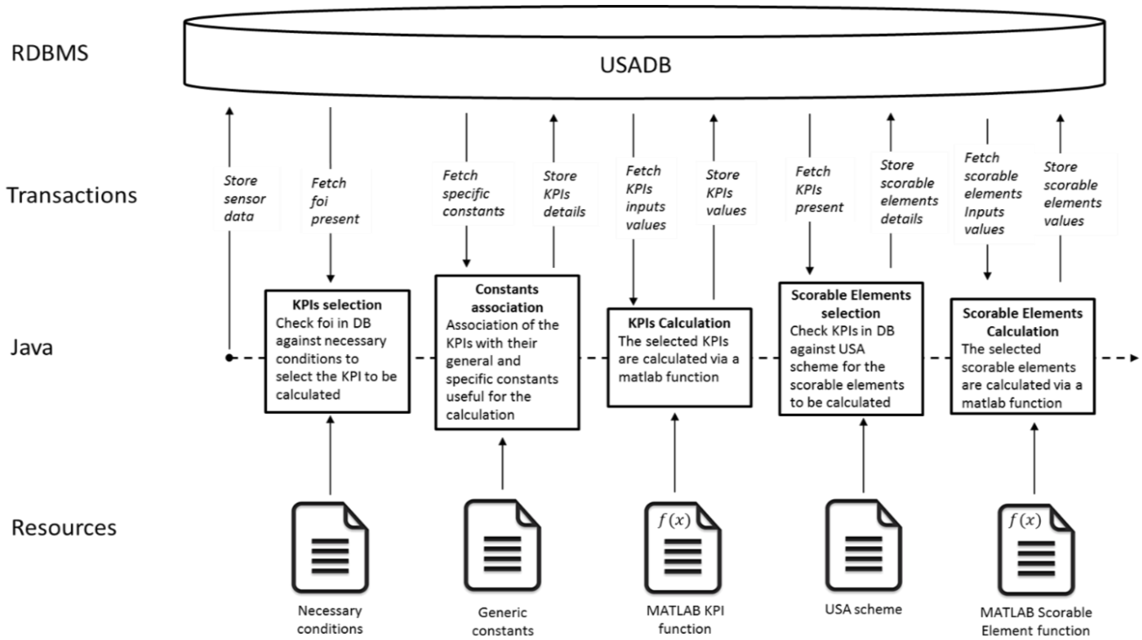


Figure 6-4 KPI Calculation Procedure

Following the procedure in Figure 6-4, the system will first fetch the features of interest (FOI) present into the RDBMS. In the present framework, the features of interest identifiers related to measurements are labelled in a hierarchical structure. For instance, the energy sources are labelled such as:

EbbwVale/EnergySource/Gaz/Boiler_Heat_EnergyProduced
EbbwVale/EnergySource/Renewable/Biomass_Heat_EnergyProduced
EbbwVale/EnergySource/Renewable/Biomass_Waste_VolumeConsumed
EbbwVale/EnergySource/Renewable/CHP_Electricity_EnergyProduced
EbbwVale/EnergySource/Renewable/CHP_Heat_EnergyProduced

Such labelling is essential to uniquely identify each feature of interest within the same urban system and to ease retrieval of information for the KPI calculator.

The FOI identifiers are then compared against necessary conditions for the KPIs calculation. Those conditions include the features of interest required as inputs for the KPIs calculation. Table 6-5 gives the necessary conditions for the KPIs to be investigated. The "%" symbol represents a wild card allowing the system to fetch all the features of interest corresponding to the string pattern. In this way, the system will only consider the appropriate KPIs to be calculated.

Table 6-5 KPIs calculation necessary conditions

KPI	Necessary conditions
AirQualityLevelIndicator	%/EnergySource/%_Electricity_EnergyProduced; %/EnergySource/%_Heat_EnergyProduced
CarbonEmissionsIndicator	%Electricity_EnergyDemand; %Heat_EnergyDemand
ComplaintRegardingThermalComfortIndicator	%Air_IndoorTemperature; %Air_RadianTemperature; %Air_IndoorRelativeHumidity
CostEffectiveDevelopmentIndicator	%Electricity_EnergyDemand; %Heat_EnergyDemand
EfficientHVACIndicator	%/HeatPump%_%_SupplyTemperature; %/HeatPump%_%_ReturnTemperature; %/HeatPump%_%_EnergyDemand
ElectricalLossesIndicator	%Electricity_EnergyDemand; %Electricity_EnergyProduced
ElectricityFromRenewableSourcesIndicator	%/Renewable/%_Electricity_EnergyProduced; %Electricity_EnergyDemand
GHGEmissionsIndicator	%Electricity_EnergyDemand; %Heat_EnergyDemand
HeatingFromRenewableSourcesIndicator	%/Renewable/%_Heat_EnergyProduced; %Heat_EnergyDemand
HeatLossesIndicator	%Heat_EnergyDemand; %Heat_EnergyProduced
OnSiteGenerationIndicator	%Electricity_EnergyDemand; %Electricity_EnergyProduced; %Heat_EnergyDemand; %Heat_EnergyProduced
OrganicWasteForEnergyGenerationIndicator	%/EnergySource/%_Waste_VolumeConsumed; %_Waste_VolumeProduced
ThermalGradientDifferencesIndicator	%Air_OutdoorTemperature; %Rural%Air_OutdoorTemperature
TimeoutThermalComfortLevelsIndicator	%Air_IndoorTemperature
TotalEnergyDemandIndicator	%Electricity_EnergyDemand; %Heat_EnergyDemand

Those features of interest (and their measurements) representing a first set of inputs for calculation, the system will then retrieve the required constants to complete the inputs set. There are two types of constant: the generic constants (e.g. basic clothing insulation, the rate of mechanical work, unit conversion …) and constants specific to a feature of interest (e.g. energy source co2 emission rate, ideal indoor temperature range, gas cost …). Ideally, the specific constants have been stored with the sensor meta-data and can, therefore, be retrieved from the RDB while generic constants are stored into a static resource file.

The KPIs details, including their inputs and labels, are then stored into the RDBMS. Inputs are fetched and sent to a third party MATLAB function that will calculate half hourly KPIs' values. Subsequently, they are stored and linked to the corresponding KPI within the database.

Finally, the remaining scorable elements (themes, subthemes and criteria) scores are determined and stored in the same fashion, checking which KPIs are present and calculating accordingly.

6.3.3.2 *KPI calculation*

The KPIs selected in Section 4.4 have been calculated half hourly using MATLAB. The calculation method uses the inputs presented in Table 6-5 to estimate the KPIs. Table 6-6 details the equations used for the KPIs calculation.

Table 6-6 KPIs calculation equations

AirQualityLevelIndicator	If $\sum EnergyProduced * NoxEmissionsRate \geq NoxLevelLimit$ $Measurement = TRUE$ $= \frac{\sum_{24h} Measurement_{TRUE}}{\sum_{24h} Measurement}$
CarbonEmissionsIndicator	$= \sum EnergyProduced * SourceCO2EmissionRate + (TotalEnergyDemand - TotalEnergyProduced) * NationalGridCO2EmissionsRate$
ComplaintRegardingThermalComfortIndicator	Predicted Percentage of Dissatisfied (PPD) calculation [549], [550]
CostEffectiveDevelopmentIndicator	$= \sum EnergyProduced * SourceCost + (TotalEnergyDemand - TotalEnergyProduced) * NationalGridCosts$
EfficientHVACIndicator	COP, EER, SEER calculation
ElectricalLossesIndicator	If $ElectricityGrid \leq 0$ $= \frac{ElectricityProduced - ElectricityDemand + ElectricityGri}{ElectricityProduced}$ Else $= \frac{ElectricityProduced - ElectricityDemand + ElectricityGri}{ElectricityProduced + ElectricityGrid}$
ElectricityFromRenewableSourcesIndicator	$= \frac{\sum RenewableElectricitySources}{TotalElectricityProduced}$

EquippedBuildingsIndicator	$= \frac{\sum Building_{withICTs}}{\sum Building}$
GHGEmissionsIndicator	$= \sum EnergyProduced * SourceGHGEmissionRate + (TotalEnergyDemand - TotalEnergyProduced) * NationalGridGHGEmissionsRate$
HeatingFromRenewableSourcesIndicator	$= \frac{\sum RenewableHeatSources}{TotalHeatDemand}$
HeatLossesIndicator	$= \frac{HeatProduced - HeatDemand}{HeatProduced}$
MonitoringSystemsIndicator	$= \frac{\sum MonitoringSytems}{\sum Building}$
OnSiteGenerationIndicator	$= \frac{TotalEnergyProduced}{TotalEnergyDemand}$
OrganicWasteForEnergyGenerationIndicator	$= \frac{TotalWasteProduced}{BiomassWasteConsumed}$
ThermalGradientDifferencesIndicator	$= OutdoorTemperature_{urbansystem} - OutdoorTemperature_{nearestRuralZone}$
TimeoutThermalComfortLevelsIndicator	If $IndoorTempereature \leq LowerLimit\ or \geq UpperLimit$ $Measurement = TRUE$ $= \frac{\sum_{24h} Measurement_{TRUE}}{\sum_{24h} Measurement}$
TotalEnergyDemandIndicator	$= \sum BuildingEnergyDemand$

Note that those calculation methods have been developed on the basis of the available sensors data. Ideally, some KPIs would be determined based on more accurate measurements. For instance, air quality levels and emissions can be measured using specific sensors and procedures [567], [641] instead of simply being correlated to energy usage.

6.3.3.3 *KPI forecast*

KPI prediction models have been developed using ANN in order to estimate the next 24h to 7 days values. The ANNs have been set up using MATLAB and are composed of 60 hidden neurons layers. Inputs are designed in a similar fashion for each KPI including:

- Previous day value at the same time;
- Previous week value at the same time;
- Previous 24h values median;
- Previous 24h values median deviation;

- Hour of the predicted value;
- Weekday of the predicted value;
- Logical value the predicted value being during a weekend or not.

All the available half hourly KPIs values from September 17 2015 00:00 to October 28 2015 23:30 have been used to train the models. Prior to the training, the inputs have been pre-processed in order to smooth their trends. Outliers have been removed using a moving median method with a moving window of 6 hours.

70% of the data have been used for training while 15% used for validation and 15% for testing. The training has been done using the Levenberg-Marquardt backpropagation algorithm. Figure 6-5 and Figure 6-6 give an overview of the prediction performance for *CarbonEmissionsIndicator* and *TotalEnergyDemandIndicator* respectively.

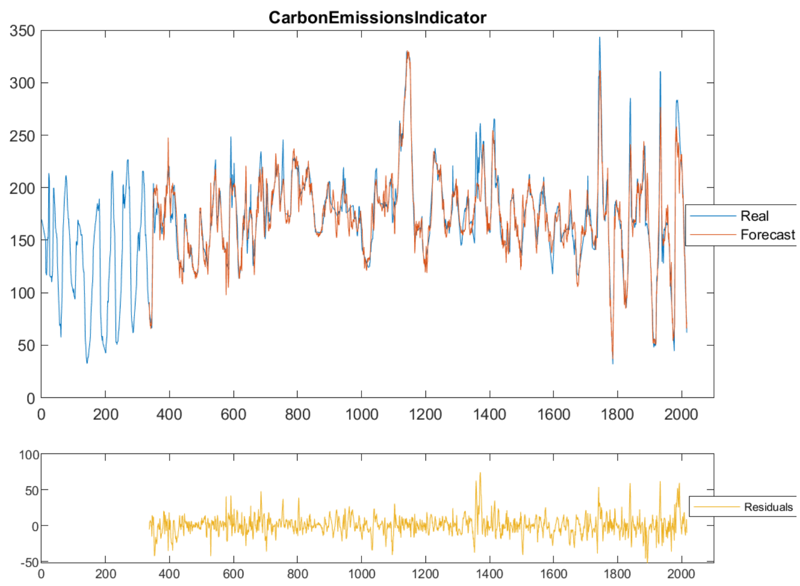


Figure 6-5 Carbon Emissions Indicator Forecast

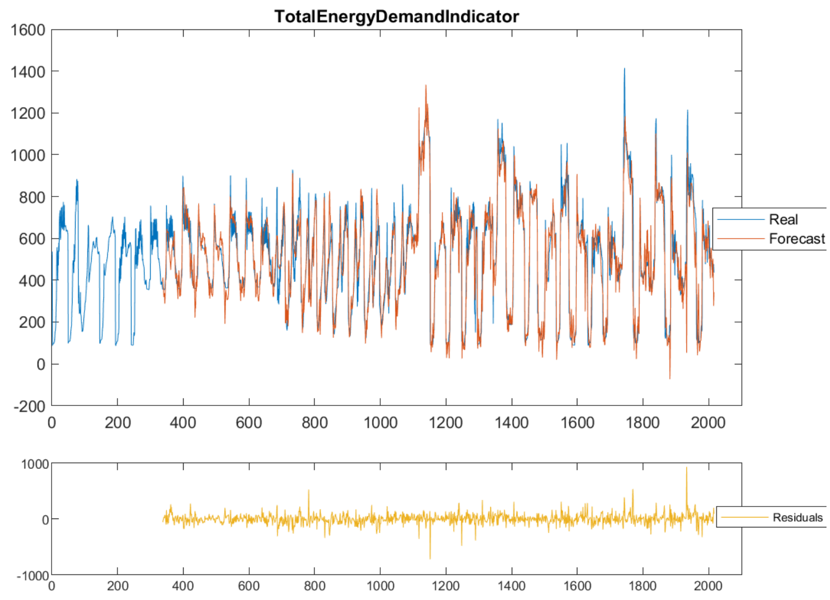


Figure 6-6 Total energy demand Indicator Forecast

Table 6-7 gives the overall performance of the forecasts for each KPIs. For each KPI, the median, the median absolute deviation (MAD) and the forecasts root mean squared errors (RMSE) are given. The forecasts show good performances apart from *HeatingFromRenewableSourcesIndicator* and *OnSiteGenerationIndicator*. The reason for that is the high frequency (half hourly) of measurement and calculation that make the KPI more volatile. Indeed, production schedule and demand are most likely asynchronous which gives sharp peaks on a high frequency. Therefore, tests have been done on hourly to every 24h aggregated data. The aggregation smooths the trends which improve the forecasts accuracy. The forecasts have reached good performance with an RMSE of 14.263 from a 6h-on aggregation for *OnSiteGenerationIndicator* and an RMSE of 14.274 from 10h-on aggregation for *HeatingFromRenewableSourcesIndicator*.

Table 6-7 KPIs 24h forecasts performances

	KPI Median	KPI MAD	Forecast RMSE
AirQualityLevelIndicator (%)	0	0	0
CarbonEmissionsIndicator (kgCO2e)	169.985	23.925	9.712
ComplaintRegardingThermalComfortIndicator (%)	13.43	4.83	2.806
ElectricalLossesIndicator (kWh)	2.155	2.155	0.549
ElectricityFromRenewableSourcesIndicator (%)	73.945	73.825	12.969
GHGEmissionsIndicator (kgCO2e)	170.36	23.985	9.926
HeatingFromRenewableSourcesIndicator (%)	112.09	49.76	241.14
HeatLossesIndicator (kWh)	34.04	27.745	17.63
OnSiteGenerationIndicator (%)	123.385	36.93	92.124

ThermalGradientDifferencesIndicator (°C)	0.9	0.32	0.004
TimeoutThermalComfortLevelsIndicator (%)	11.11	11.11	6.855
TotalEnergyDemandIndicator (kWh)	521.57	148.315	95.835

In accordance with the results of the DELPHI methods in Section 4.4.6, CostEffectiveDevelopmentIndicator and OrganicWasteForEnergyGenerationIndicator have been predicted with a one-week horizon. The same ANN has been set up with inputs being:

- Previous week value at the same time;
- Previous 168h values median;
- Previous 168h values median deviation;
- Hour of the predicted value;
- Weekday of the predicted value;
- Logical value the predicted value being during a weekend or not.

Table 6-8 gives the overall performance of the forecasts.

Table 6-8 KPIs 7 days forecasts performances

	KPI Median	KPI MAD	Forecast RMSE
CostEffectiveDevelopmentIndicator (£)	32.14	7.74	2.802
OrganicWasteForEnergyGenerationIndicator (%)	82.105	82.105	21.094

6.4 TRIPLE STORE FOR ONTOLOGY STORAGE

With RDF graph gaining importance, the development of technologies that can store them has followed. Those databases are called triple-stores or RDF stores and enable storage of triples for both concepts (TBox) and assertions (ABox) [278]. Similarly to RDBMS that can be queried and manipulated using the SQL language, a triple store can be queried using the SPARQL language.

Triple stores holds a number of advantages[485]:

- Schema flexibility and simplicity allows an easy data load;
- Ease to query over distributed sources;
- Standardised which ease migration from a triple store to another;
- Provenance tracking is easier which ensure data quality;

- Query expressivity make it easier to retrieve information, unlike RDBMS where one needs to know the tables joints;
- Reasoning allows the discovery of new information inferred from the existing ones;
- Economic use of resources reducing overall costs;
- Internally stored information can be enhanced by levering the linked open data cloud available.

The distinction between information stored in the triple store and information stored in the RDBMS will depend on their volume and impact the reasoning capacity. Here, “finite” data which are the ones used for simple objects description, are distinguished with “infinite” data which are typically time series. Consequently, the triple store will hold the USA ontology concepts and roles (TBox) as well as certain individual assertions (ABox) such as BIM and cityGML entities, unit of measure or time-related information. Only sensors data and metadata will be stored in the MySQL database.

In this research, the USA ontology has been developed using the Protégé framework [642] and is therefore stored in RDF files accessed by the software. Protégé is an open-source platform that provides an interface to build and query ontologies. The tool benefits from a growing community of users and is strengthened by a broad variety of plug-in that extended its capabilities such as DL queries, ONTOP (discussed in the next section), SWRL rules (Semantic Web Rule Language) or graph visualisation for instance.

Storing BIM models in triple stores has already been investigated during the development of the ifcOWL ontology [335], [337], [344], [618]. As mentioned in Section 2.3.3.2., a BIM model can be instantiated in RDF simply from scratch with the appropriate tools such as Protégé, Jena or the OWL API for Java, task that can be however time-consuming if one wants a high level of detail. In that regards, an automatic converting tool has been developed to convert IFC files into RDF graph [337]–[342]. In this study, an IFC-to-RDF converter has been used to convert BIM models of the Ebbw Vale site using a free tool available online [343]. This JAVA-based tool has been developed by the research team responsible for the ifcOWL standard. The first attempt to convert the Learning Zone BIM model has resulted in the creation of a 545MB RDF file.

In term of efficiency and performance, the size of the ifcOWL graphs matters. Efforts can be done in constructing simpler RDF graphs from the BIM model. Indeed, ifcOWL graphs are complex due to their direct interpretation of the IFC EXPRESS schema. Consequently, some IFC data constructs present in ifcOWL are rarely used for web semantic and are therefore unnecessarily modelled. This applies for instance to the geometries and the detail description of Cartesian points, polylines or segment descriptions for example. Figure 6-7 presents a solution to this issue where RDF graphs are restructured into a set of function-related RDF graphs.

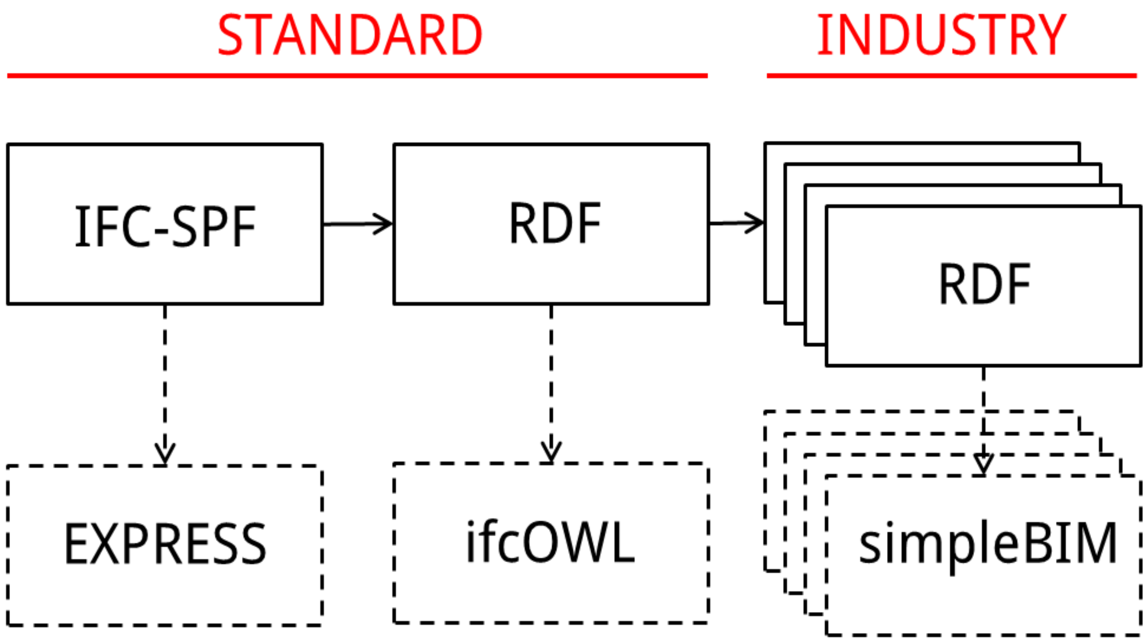


Figure 6-7 simpleBIM conversion [38]

This can lead to the generation of alternative simpler graphs (simpleBIM) linked to the original IFC [618], [643]. Ultimately, a tool has been developed that takes as input an IFC files, removes a number of entities and send back a restructured simpler ifcOWL graph of reasonable size. The procedure removes the following entities from the IFC-RDF conversion:

- presentation resource
- presentation appearance resource
- presentation definition resource
- profile resource
- representation resource
- topology resource
- geometry resource
- geometric model resource
- geometric constraint resource.

Such a tool is however not publicly available as it is still under development.

The author has not found the equivalent of the IFC-to-RDF converter that would convert a cityGML model into RDF OWL format. Such conversion is mentioned in Hor's paper [266] without being implemented. However, Hor opens on the possibility to convert GML format, which is the based schema for cityGML. An interesting lead is the use of The RDF Mapping language (RML), a mapping language that allows mapping data in heterogeneous structures and serializations to the RDF data model as shown in Figure 6-8 [644]–[646]. Indeed, the GML format (and in turn cityGML) is XML-based format which can be mapped to RDF via RML mapping rules. If the conversion of GML formatted GIS have been the subject to many studies [647]–[649], nothing has been found concerning cityGML files conversion via this method.

Overall, the immaturity of the conversion tools has led the author to instantiate the model from scratch with a low level of detail as it is less time-consuming.

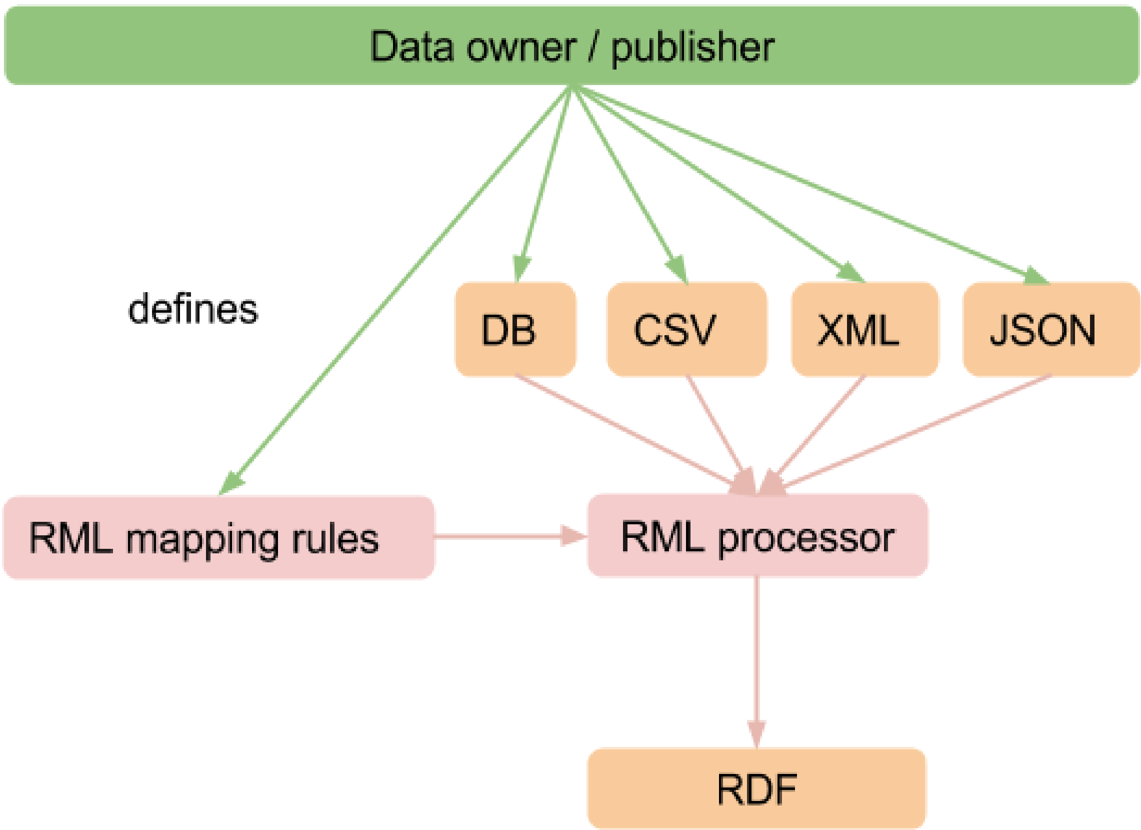


Figure 6-8 RML overview [644]

6.5 ONTOP: FROM RELATIONAL DATABASE TO LINKED DATA

ONTOP [624] is the core element of the OBDA that allows the linkage between sensor data and metadata stored in the MySQL relational database, and the concepts, roles, BIM and cityGML assertions present in the triple store. This JAVA-based framework allows SPARQL queries over virtual RDF graphs generated through RDBMS mappings. Those virtual graphs are created on-the-fly which confers high performance of the system over very large data sets and ontologies. The tool benefits from an active development and a growing community [647]. In terms of programming, it is very advantageous as it presents itself in different distribution settings such as a set of JAVA libraries/scripts to use with the OWLAPI, a plugin for Protégé or a jetty/tomcat bundle for the deployment of web apps.

Figure 6-9 gives an overview of the ONTOP system structure. The system can be divided into two distinct parts: the input files that one provides to the system and the Quest engine which is the core of the system.

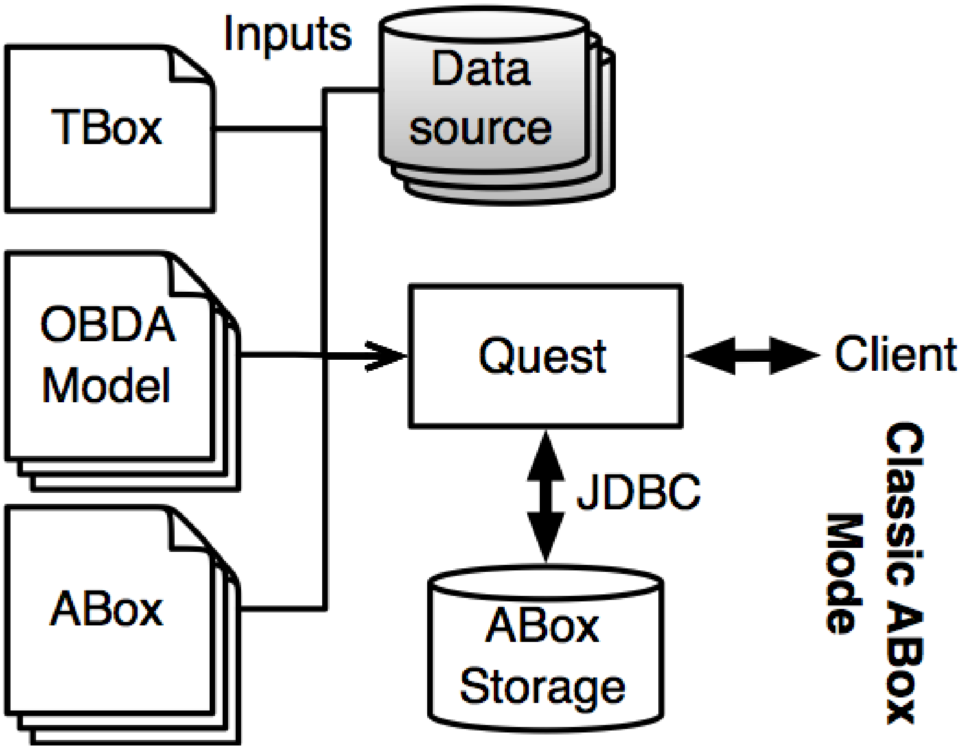


Figure 6-9 ONTOP structure

6.5.1 ONTOP Inputs

The TBox and ABox are both parts of the OWL ontology inputed in the system. ONTOP will process the OWL graph and construct a virtual ABox complementary of the assertions already present in the actual ABox. In this way, one can connect already asserted individuals (e.g. a BIM model) to data stored in the RDBMS. Note that the QUEST engine will RDFS/OWL2QL interpret the input ontology. The OWL 2 QL entailment, already presented in Section 2.3.2.2, supports the following axioms:

- subclass assertions
- disjointness (DisjointClasses)
- disjoint properties (DisjointObjectProperties and DisjointDataProperties)
- symmetric properties (SymmetricObjectProperty)
- DifferentIndividuals

The RDFS entailment is non-standard with only a limited amount of axioms supported:

- rdfs:type
- rdfs:subClassOf
- rdfs:subPropertyOf
- rdfs:domain
- rdfs:range

THE OBDA model is composed of 3 types of documents: the Properties document that will hold the JDBC connection parameters, the mapping files that list the set of mappings between SQL queries and RDF axioms, and the SPARQL file where “classic” SPARQL queries are formulated.

6.5.1.1 *ONTOP JDBC connection settings*

The ONTOP system uses a JDBC driver to connect to the MySQL database and perform SQL queries. Figure 6-10 gives a view of the ONTOP JDBC settings window where the user must provide connection parameter namely the connection URL, Username, Password and JDBC driver class.

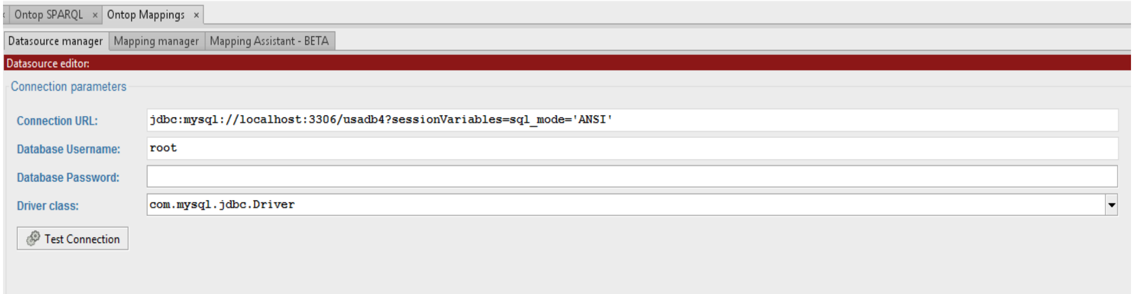


Figure 6-10 ONTOP JDBC settings window

The type of driver used depends on the type of targeted RDBMS and it is up to the user to provide the appropriate driver when setting up the OBDA. ONTOP supports the integration of the MySQL, PostgreSQL, H2, DB2, Microsoft SQL Server, Oracle, Teiid and MonetDB databases and their respective JDBC drivers.

6.5.1.2 *ONTOP mappings*

ONTOP features its own mapping system where a user provides an SQL source query and an OWL assertion axiom (the target). Figure 6-11 gives an example that illustrates the definition of ONTOP mappings. It shows how individuals stored in the RDBMS (in black) fit in classes and connect to each other or to individuals stored in the triple store via some object property or data property.

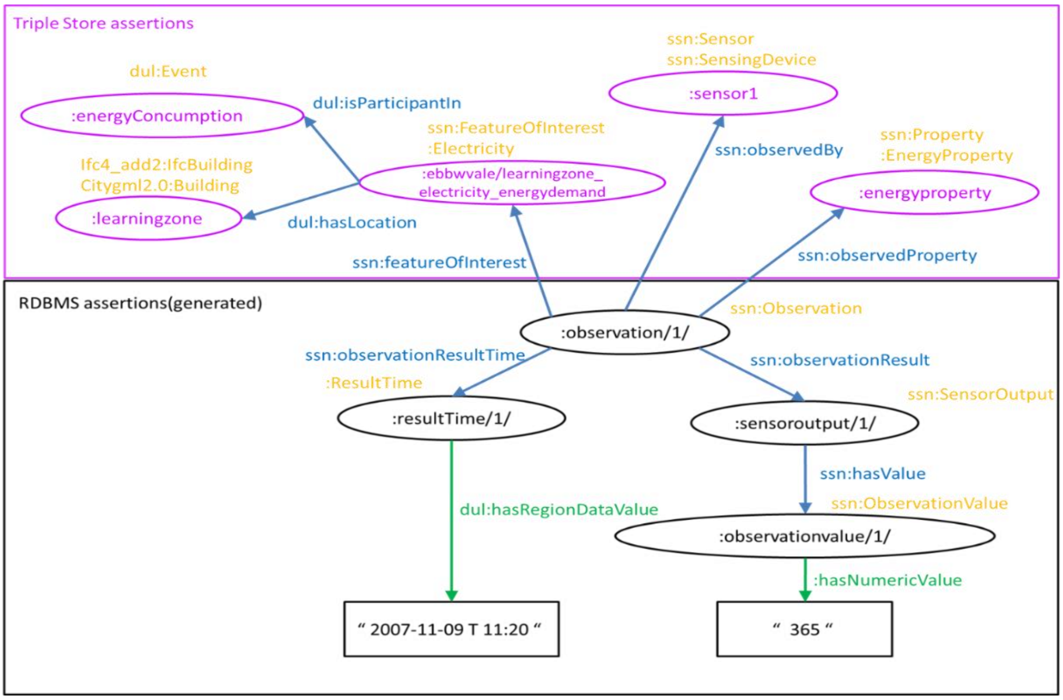


Figure 6-11 USA ontology assertions example

Based on the example given in Figure 6-11, Figure 6-12 illustrates how the mapping editor allows the user to provide OWL assertion axioms and SQL query centred on *ssn:Observation.* The class is instantiated with individuals defined by the *:observation/{observationId}/* URI where the part in between *{}* will be asserted with the results of the SQL query corresponding to *observationId* shown at the bottom of the window. The features of interest, properties and sensors will, therefore, be asserted with their unique identifier label stored in the RDBMS. It is essential for those unique identifiers to be identical to the ones present in the triple store in order to ensure a consistent linkage. Indeed, two instances with the same URI will be considered as one by the ONTOP system.

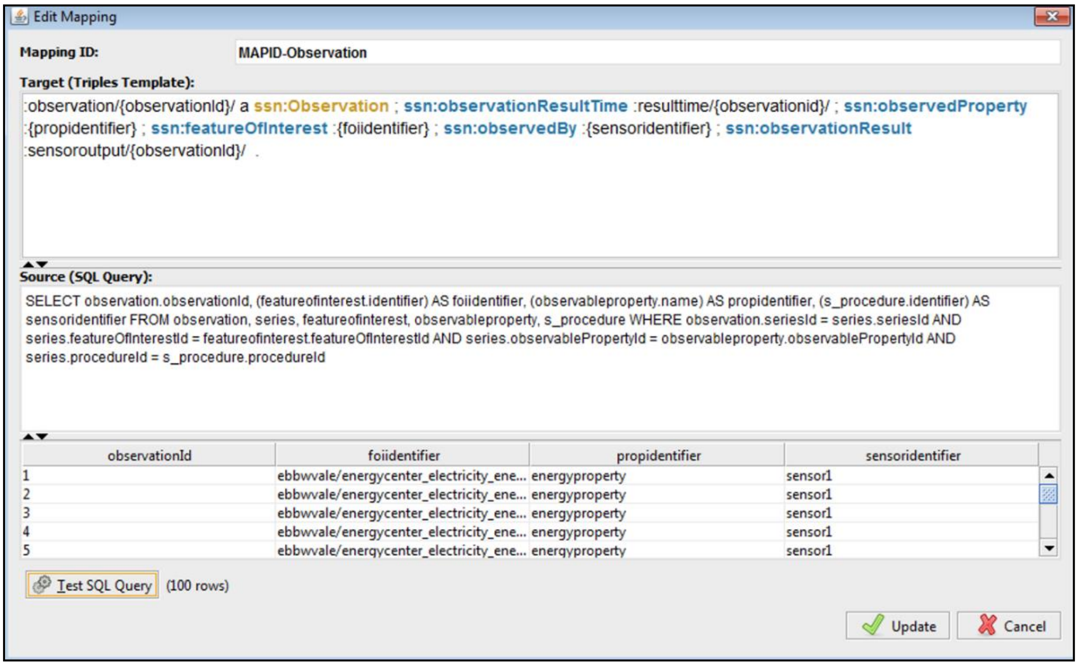


Figure 6-12 ONTOP Observation Mapping

In the same way, Figure 6-13 shows how the mapping of a *SensorOutput* and its relationship with *ObservationValue* can be designed.

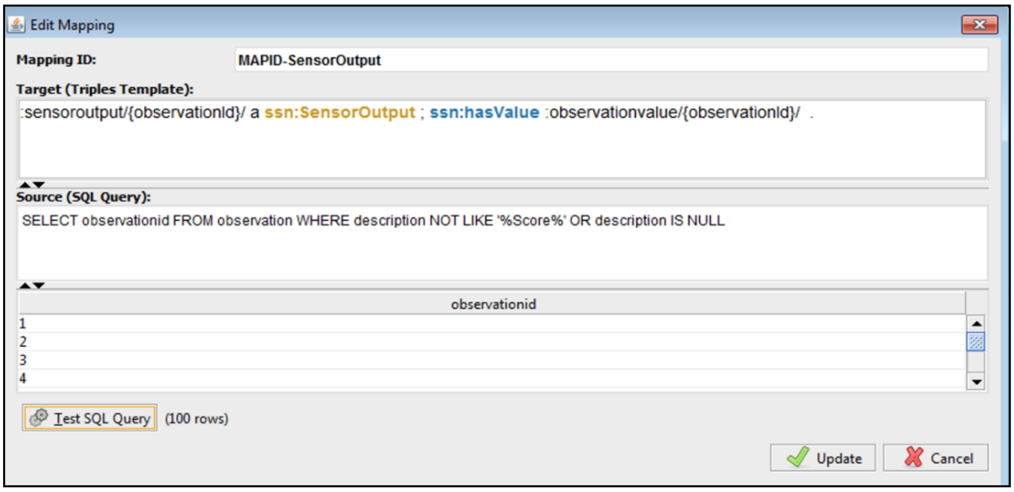


Figure 6-13 ONTOP SensorOutput mapping

Figure 6-14 gives an insight of how data properties can be asserted using ONTOP mappings. *ResultTime* is given some dateTime values taken from the MySQL DB while *ObservationValue* is instantiated with decimal values. It is important for the formats defined in the target axioms (e.g. xsd:dateTime, xsd:decimal) to be compatible with the ones defined in the RDBMS as doing otherwise could result in errors during the process.

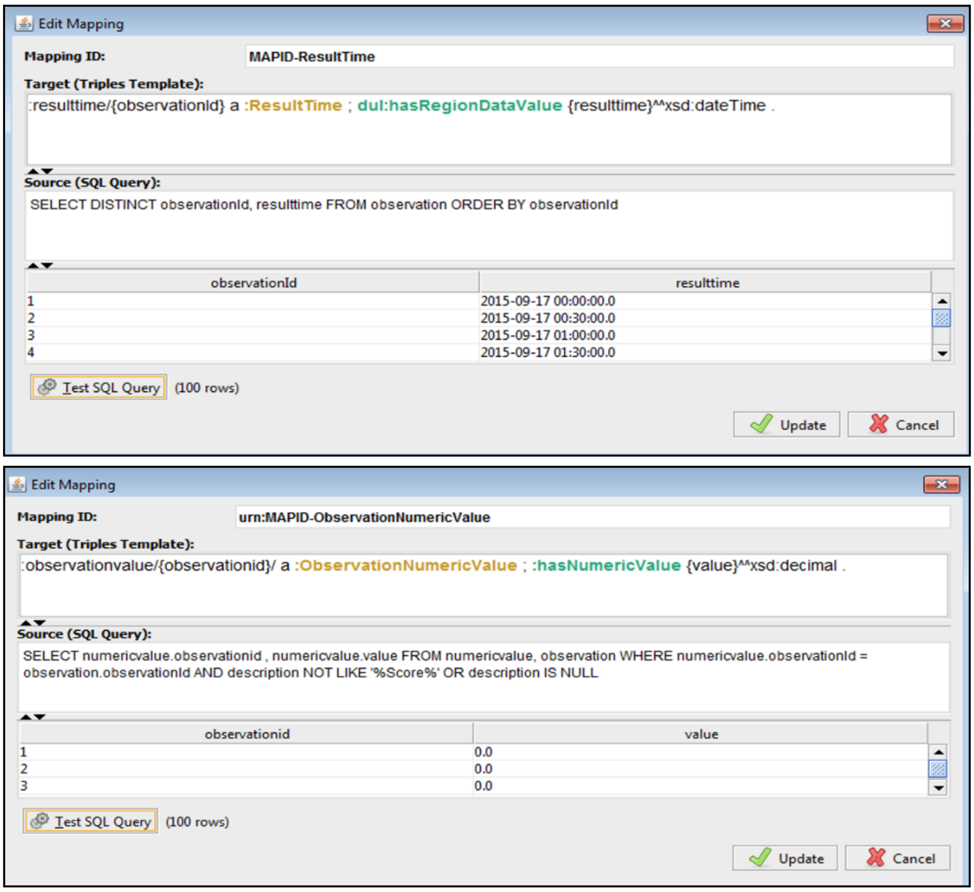


Figure 6-14 ONTOP ResultTIme and IndicatorNumericValue mapping

Finally, Figure 6-15 shows a mapping where individuals stored in the RDBMS (*:observation/{observationid}/)* are linked to an individual stored in the triple store (:ebbwvale), a member of the *UrbanSystem* class. The difference with the mapping presented in Figure 6-12 where the piece of information that linked an observation to a feature of interest was stored in the RDBMS, is that here the piece of information is uniquely present in the mapping.

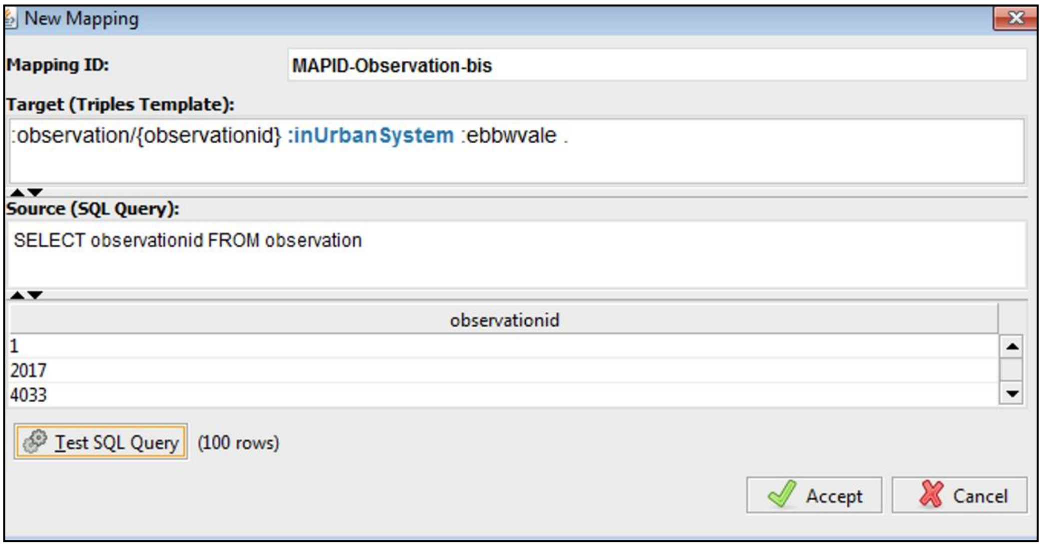


Figure 6-15 ONTOP urban system mapping

Overall, these five mappings cover the example given in Figure 6-11. When developing mappings, one must be careful to be consistent in the labels given to the individuals across the target statements as the reasoner consider them to be equivalent.

6.5.1.3 *ONTOP queries*

The queries are done in a similar fashion of the SPARQL 1.1 queries. However, some restrictions exist due to the OWL2QL entailment. ASK, CONSTRUCT and DESCRIBE queries are available as well as REGEX SPARQL filter that defines search patterns. Results are given in a tabular way and can be exported into CSV file format. Figure 6-16 shows the ONTOP SPARQL endpoint with the query editor window and the answer table window. In this example, one wants to retrieve all the observations with their values, time and unit associated. The query fetches 84,626 records in 1.478 seconds.

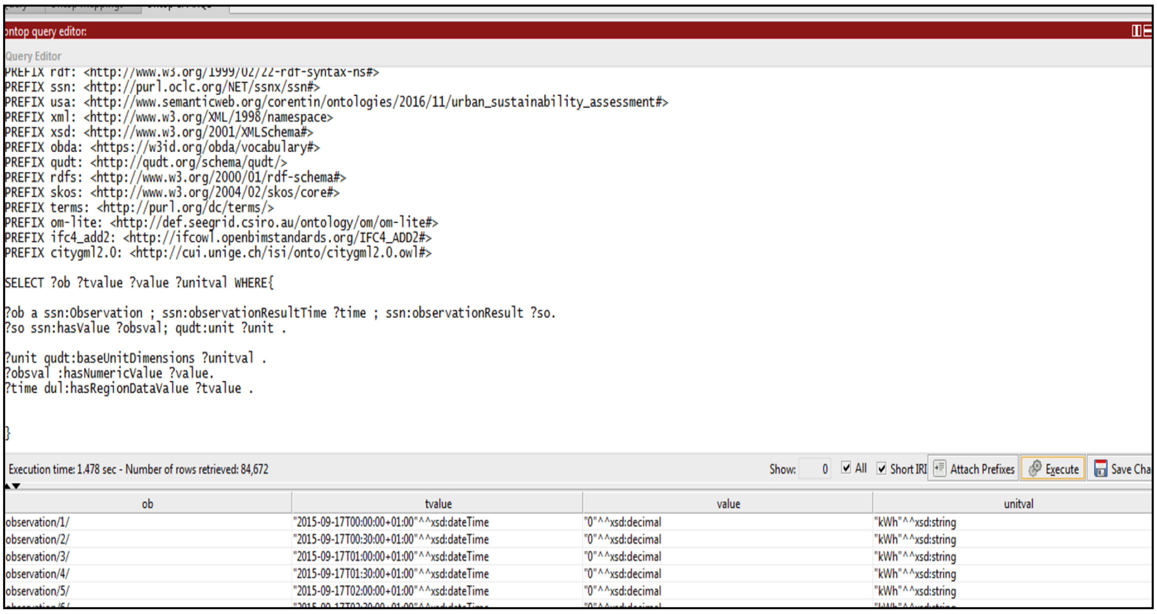


Figure 6-16 ONTOP SPARQL Query

Query 1 demonstrates how the REGEX function could be used by the user to filter the result containing a certain substring.

Query 1. Retrieve the feature of interest that contains the substring "LearningZone"

SPARQL query	Output (5 records in 0.028 sec)

SELECT * **WHERE**{
 	
?foi a ssn:FeatureOfInterest ; dul:hasDataValue ?foiname.
FILTER regex(?foiname,"Learning Zone")
 	
}
	:ebbwvale/learningzone_air_indoorrelativehumidity "EbbwVale/LearningZone_Air_IndoorRelativeHumidity"^^xsd:string
	
	:ebbwvale/learningzone_air_indoortemperature "EbbwVale/LearningZone_Air_IndoorTemperature"^^xsd:string
	
	:ebbwvale/learningzone_air_radiantemperature "EbbwVale/LearningZone_Air_RadianTemperature"^^xsd:string
	
	:ebbwvale/learningzone_electricity_energydemand "EbbwVale/LearningZone_Electricity_EnergyDemand"^^xsd:string
	
	:ebbwvale/learningzone_heat_energydemand "EbbwVale/LearningZone_Heat_EnergyDemand"^^xsd:string

Query 2 demonstrates the inclusion of the BIND construct that allows assigning the result of an expression to a variable. In the present case, the string "`EbbwVale/LearningZone_`" is removed from the feature of interest name.

Query 2. Using BIND to create a new variable

SPARQL query	Output (5 records in 0.1 sec)
`SELECT ?foi ?newfoiname WHERE{ ?foi a ssn:FeatureOfInterest ; dul:hasDataValue ?foiname. FILTER regex(?foiname,"LearningZone"). BIND( REPLACE( ?foiname,"EbbwVale/LearningZone_", "") AS ?newfoiname ). }`	:ebbwvale/learningzone_air_indoorrelativehumidity "Air_IndoorRelativeHumidity"^^xsd:string :ebbwvale/learningzone_air_indoortemperature "Air_IndoorTemperature"^^xsd:string :ebbwvale/learningzone_air_radiantemperature "Air_RadianTemperature"^^xsd:string :ebbwvale/learningzone_electricity_energydemand "Electricity_EnergyDemand"^^xsd:string :ebbwvale/learningzone_heat_energydemand "Heat_EnergyDemand"^^xsd:string

Query 3 uses the VALUES construct in order to filter the property name corresponding to "*EnergyProperty*" and "*TemperatureProperty*". With such feature, one can input a list of variables'name that he or she is interested in. The construct ORDER BY is equally used in this example to order the result set following the observed property ID.

Query 3. Retrieve EnergyProperty and TemperatureProperty using VALUES construct

SPARQL query	Output (2 records in 0.019 sec)
`SELECT * WHERE{ ?prop a ssn:Property; dul:hasDataValue ?propname. VALUES ?propname {"EnergyProperty" "TemperatureProperty"}. }ORDER BY ?prop`	:energyproperty "energyproperty"^^xsd:string :temperatureproperty "temperatureproperty"^^xsd:string

Query 4 demonstrates the use of UNION in order to retrieve both values of *GHGEmissionsIndicator* and *TotalEnergyIndicator*.

Query 4. Using UNION to combine graph patterns

SPARQL query	Output (4032 records in 1.017 sec)

`SELECT DISTINCT * WHERE{`
 	
`{?so rdf:type usa:GHGEmissionsIndicator.}`
`UNION`
`{?so rdf:type usa:TotalEnergyDemandIndicator.}`
 	
`?so ssn:hasValue [usa:hasNumericValue ?value].`
 	
`}`
	`:sensoroutput/108865/ "537"^^xsd:decimal`
	
	`:sensoroutput/108867/ "87"^^xsd:decimal`
	
	`:sensoroutput/108869/ "95"^^xsd:decimal`
	
	`:sensoroutput/108871/ "96"^^xsd:decimal`
	`...`

Other functions such as BOUND, CONCAT, isIRI, isBlank, str, SUBSTR, UCASE, LCASE, CONTAINS, STRBEFORE, STRAFTER, ABS, ROUND, FLOOR, CEIL, RAND, now, year, month, day, hours, tz etc are supported.

6.5.2 ONTOP Quest

The QUEST engine is at the core of the ONTOP system and allows the translation of SPARQL queries into SQL queries. Figure 6-17, taken from Calvanese et al. [650], shows the QUEST engine structure in more details. The model is split in two distinct part: the off-line and the on-line stages.

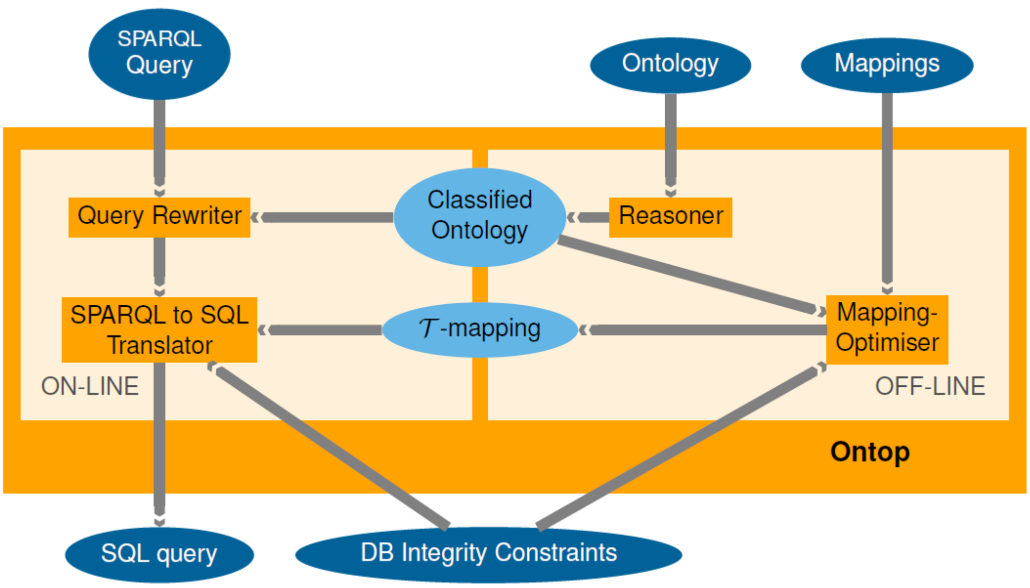


Figure 6-17 ONTOP QUEST structure [650]

Starting with the off-line stage, the ontology and mapping are first introduced by the user.

Example 1.1. Ontology sample

```
:HeatFromRenewableSources  rdfs:subClassOf  :Heat .
              :HeatLosses  rdfs:subClassOf  :Heat .
                    :Heat  rdfs:subClassOf  :EnergyObject .
          dul:hasLocation  rdfs:range :Position .
```

Firstly, the ontology reasoner will run resulting in the creation of a classified ontology with the complete hierarchy of properties and classes and all inferred axioms. This classified ontology is stored in memory.

Example 1.2. Inferred axioms after reasoning

```
:HeatFromRenewableSources  rdfs:subClassOf  :EnergyObject .
              :HeatLosses  rdfs:subClassOf  :EnergyObject .
```

Then, using the classified ontology and the mapping introduced as inputs, a mapping optimiser will compose new mappings, the T-mappings that catches hidden information from the inferred axioms.

Example 1.3. Inferred axioms T-mappings

```
:foi/{featureOfInterestId}/  rdf:type  :EnergyObject .
   ← SELECT featureofinterest.featureOfInterestId FROM featureofinterest
      WHERE featureofinterest.foiname = 'HeatFromRenewableSources'
:foi/{featureOfInterestId}/  rdf:type  :EnergyObject .
   ← SELECT featureofinterest.featureOfInterestId FROM featureofinterest
      WHERE featureofinterest.foiname = 'HeatLosses'
```

The T-mappings are then optimised via disjunction (OR) and interval expressions and by applying Semantic Query Optimization techniques such as removing redundant self-joints and trivial conditions for instance [650].

Example 1.4. Optimised T-mapping

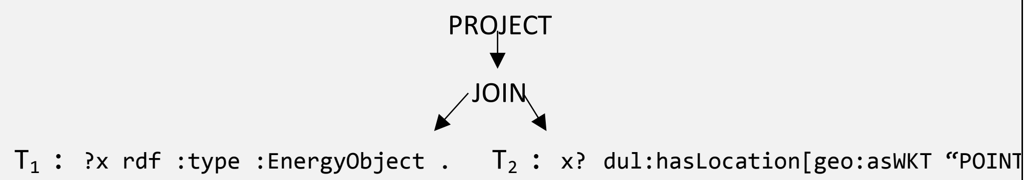
```
:foi/{featureOfInterestId}/  rdf:type  :EnergyObject .
   ← SELECT featureofinterest.featureOfInterestId FROM featureofinterest
      WHERE featureofinterest.foiname = 'HeatFromRenewableSources' OR
       featureofinterest.foiname = 'HeatLosses'
```

The creation and optimisation of the T-mappings can be expensive in term of computing. However, they are constructed once during the off-line stage and can be accessed and used subsequently during the on-line stage as a basis for the translation of SPARQL queries into SQL.

During the on-line stage, the user inputs a SPARQL query that will be re-written in a fashion that helps the translation. Indeed, the query rewriter will create a tree of objects that corresponds to the SPARQL algebra expression.

Example 1.5. SPARQL query and corresponding Tree Object

```
SELECT ?foi WHERE
       {
       ?foi rdf:type :EnergyObject ; dul:hasLocation [geo:asWKT “POINT(-3.2039
       51.7766)”^^geo:wktLiteral] .
       }
```



An algorithm goes through the entire tree hierarchy in a bottom-up fashion, starting from the leaves and translating each SPARQL nodes into its corresponding SQL statement as it iterates through them.

Example 1.6. Corresponding SQL query

```
SELECT Q1.x FROM
      ((SELECT featureofinterest.featureOfInterestId AS x FROM featureofinterest
        WHERE featureofinterest.foiname = 'HeatFromRenewableSources' OR
 featureofinterest.foiname = 'HeatLosses') Q1
       JOIN
       (SELECT featureofinterest.featureOfInterestId AS x FROM featureofinterest
        WHERE featureofinterest.geom = 'POINT(-3.2039 51.7766) ') Q2
       ON Q1.x = Q2.x)
```

The query is then optimised and simplified using the same Semantic Query Optimization techniques mentioned earlier in this section.

Example 1.7. Final Optimised SQL query

```
SELECT featureofinterest.featureOfInterestId  FROM featureofinterest
       WHERE featureofinterest.foiname = 'HeatFromRenewableSources'
       OR  featureofinterest.foiname = 'HeatLosses'
       AND featureofinterest.geom = 'POINT(-3.2039952278137207 51.776660268454144)'
```

This online stage of the QUEST engine is what confers the On-the-fly aspect to the ONTOP tool with a live translation of the SPARQL query and the retrieval of the corresponding individual assertions.

6.5.3 ONTOP Limits

6.5.3.1 *Entailment limitations*

ONTOP presents a series of limitations in terms of expressivity that limits its uses. Indeed, as mentioned earlier, the tool is based in OWL 2 QL, a fragment of the OWL 2 language. Found in the specification document [306], OWL 2 QL does not support the following axioms:

- *"existential quantification to a class expression or a data range (ObjectSomeValuesFrom and DataSomeValuesFrom) in the subclass position*
- *self-restriction (ObjectHasSelf)*
- *existential quantification to an individual or a literal (ObjectHasValue, DataHasValue)*
- *enumeration of individuals and literals (ObjectOneOf, DataOneOf)*
- *universal quantification to a class expression or a data range (ObjectAllValuesFrom, DataAllValuesFrom)*
- *cardinality restrictions (ObjectMaxCardinality, ObjectMinCardinality, ObjectExactCardinality, DataMaxCardinality, DataMinCardinality, DataExactCardinality)*
- *disjunction (ObjectUnionOf, DisjointUnion, and DataUnionOf)*
- *property inclusions (SubObjectPropertyOf) involving property chains*
- *functional and inverse-functional properties (FunctionalObjectProperty, InverseFunctionalObjectProperty, and FunctionalDataProperty)*
- *transitive properties (TransitiveObjectProperty)*
- *keys (HasKey)"* [306]

As a workaround, the user can develop more explicit expressions and/or mappings to cover the missing inferences normally derived from those constructs. For instance, in Example 2, the transitive object property *ssn:observes* should infer that a certain "sensor" is linked to a certain "property".

Example 2.1.

```
ssn:Sensor ssn:madeObservation ssn:Observation.
ssn:Observation ssn:observedProperty ssn:Property.

ssn:madeObservation o ssn:observedProperty rdfs:subPropertyOf ssn:observes
```

Therefore, one could assume that by creating mappings for the first 2 axioms of Example 2 and using the object property "observes", he or she could retrieve which property is linked to which sensor. However, such query is impossible in ONTOP because transitivity is not supported as shown in Query 5.1. Such issues can be overcome by developing a more complex query, requiring the use of a joint and the DISTINCT construct like in Query 5.2. However, this will considerably increase the computing power required and reasoning time. The most effective solution will be therefore to create a mapping that will catch the relationship covered by the object property *ssn:observes.*

Query 5. Retrieving sensor and property relationship

1. SPARQL query without mapping	Output

`SELECT * WHERE{`
`?sensor ssn:observes ?op .`
`}`
	No records

2. Complex SPARQL query without mapping	Output

`SELECT DISTINCT ?sensor ?op WHERE{`
`?sensor ssn:madeObservation ?ob.`
`?ob a ssn:Observation;`
`ssn:observedProperty ?op .`
`}`
	69 records in 3.113 sec

3. SPARQL query with mapping	Output

`SELECT * WHERE{`
`?sensor ssn:observes ?op .`
`}`
	69 records in 0.026 sec

Overall, mapping design should ensure that the use of DISTINCT in the SPARQL is limited. The designer should make sure that each axiom only appears once, to label the individuals with tables Primary and foreign key, to avoid the use of UNION and unnecessary joins in SQL and be consistent in the entities naming.

6.5.3.2 *SPARQL limitations*

Additionally, there is a set of unsupported function in ONTOP SPARQL, the most constraining being:

- Aggregate functions such as AVG, SUM, MIN, MAX, COUNT, SAMPLE, GROUP BY (support in development);
- Cast function (e.g. xsd:string(...), xsd:integer(...));
- Custom datatypes.

A solution has been investigated within the THERMOSS project in order to overcome the limitation w.r.t the aggregated functions. Indeed, a technical requirement of the THERMOSS OBDA implementation was the possibility to identify faulty data from time series in an easy manner. Such task requires the detection of outliers via statistical approaches including the median absolute deviation (MAD).

Query 6. Annotation example of aggregated function using SELECT

SPARQL query	Output (208 records in 4.837 sec)

SELECT * **WHERE**{
 	
#outliers
thermoss:property_60
saref:relatesToMeasurement ?data.
?data saref:hasTimestamp ?moment1;
saref:hasValue ?value1 .
 	
}
	:datas_28335
	"2018-08-
	24T13:10:00+01:00"^^xsd:dateTime
	"19.0"^^xsd:double
	
	:datas_28336
	"2018-08-
	24T13:15:00+01:00"^^xsd:dateTime
	"19.0"^^xsd:double
	
	:datas_28337
	"2018-08-
	24T13:20:00+01:00"^^xsd:dateTime
	"19.0"^^xsd:double
	
	:datas_28338
	"2018-08-
	24T13:25:00+01:00"^^xsd:dateTime
	"19.0"^^xsd:double

Query 6 gives an example of the outliers retrieval executed from the OWL API in JAVA. The system is based on annotation where one enters the function desired, here *"#outliers*", along the query to retrieve the time series investigated. Note that if the THERMOSS project uses the SAREF ontology for its OBDA as the ontology was closer to the intended application, the procedure would be consistent with any type of OBDA that can access time series.

In the example, the time series associated with the *thermoss:property_60* have been retrieved and the outliers have been calculated internally within the JAVA procedure. This allowed to pick up which triples were designated as being an outlier and to output them.

Query 7. Annotation example of aggregated function using CONSTRUCT

SPARQL query	Output (208 records in 4.842 sec)

CONSTRUCT {thermoss:property_60 thermoss:hasOutliers ?data}
WHERE{
 	
#outliers
thermoss:property_60 saref:relatesToMeasurement ?data.
?data saref:hasTimestamp ?moment1; saref:hasValue ?value1 .
 	
}
	ObjectPropertyAssertion(thermoss:hasOutliers, thermoss:property_60, thermoss:datas_28335)
	
	ObjectPropertyAssertion(thermoss:hasOutliers, thermoss:property_60, thermoss:datas_28336)
	
	ObjectPropertyAssertion(thermoss:hasOutliers, thermoss:property_60, thermoss:datas_28337)

Additionally, using the CONSTRUCT method as in Query 7 allows the user to store the outliers in the ontology as being outliers using the object property *thermoss:hasOutliers.*

6.5.3.3 *Geospatial function limitation and ONTOP-spatial*

ONTOP in its current form does not support geospatial functions, which greatly limits the use of GIS filters within the SPARQL queries. Consequently, ONTOP-Spatial, an extension of the ONTOP framework has been developed in order to cover geospatial aspects[651]–[653]. In the same fashion, ONTOP-Spatial creates virtual geospatial RDF graph on top of geospatial databases using the standard geoSPARQL ontology and geometry literals. Therefore, constructs like geo:overlaps can be used in a FILTER to know whether or not two geometry overlap each other. Ultimately, one can FILTER a set of sensors corresponding to a particular zone for instance. ONTOP-Spatial is however not integrated (yet) within the USA framework. Indeed, ONTOP-Spatial requires the usage of geospatial RDBMS such as the PostgreSQL database with PostGIS extension enabled. However, in its current configuration, the USA framework is using a MySQL database. Some effort will be necessary to migrate the current MySQL database into PostgreSQL in order to perform geoSPARQL queries, possibly when a more complete and followed-up version of ONTOP-Spatial is released.

6.6 QUERY TESTING ON THE ABOX

In his section, the USA framework will be evaluated against a set of queries on the ABox and the retrieval of actual sensors and KPIs values as well as their geolocation, their calculation procedure etc. Note that a set of queries on the TBox has already been equally evaluated as reported in Section 5.4.

The following set of queries is derived from the competency questions in Table 5-2 and Table 5-3 in section 5.2.3 where the values of indicators, their benchmarks, scores, units, recording times, locations and relationships with the urban environment are questioned.

Those requests have been done in Protégé bundled with ONTOP plugin using a desktop computer with 1TB HDD, Intel Core i7-4790 CPU 3.60GHz, 24 GB memory and Windows 7 64-bits.

Query 8 investigates the features of interest present within the Ebbw Vale urban system. Fifty eight features of interest are retrieved including the ones directly sensed such as the different heat demands, electricity demands, indoor temperatures, waste volume, as well as the ones calculated such as the GHG emitted, heat from renewables, heat losses, costs etc.

Query 8. What are the features of interest present in the urban system of Ebbw Vale?

SPARQL query	Output (58 records in 0.096 sec)

`SELECT DISTINCT * WHERE{`
 	
`#Selection of the feature of interest in ebbw vale urban system`
`?foi a ssn:FeatureOfInterest; :inUrbanSystem <http://www.semanticweb.org/corentin/ontologies/2016/11/urban_sustainability_assessment#ebbwvale>.`
 	
`}`
	:ebbwvale/energycenter_electricity_energydemand
	
	:ebbwvale/energycenter_heat_energydeman
	
	:ebbwvale/energysource/gaz/boiler_heat_energyproduced
	
	:ebbwvale/energysource/renewable/biomass_heat_energyproduced
	
	:ebbwvale/energysource/renewable/biomass_waste_volumeconsumed
	
	...

Query 9 gives an overview of features of interest can be filtered via their relation to specific urban objects. In the present case, the features of interest relating to the Learning Zone have been retrieved from the building unique ID, common with the BIM or cityGML model.

Query 9. What are the features of interest part of the learning zone building?

SPARQL query	Output (5 records in 0.007 sec)

`SELECT * WHERE{`
 	
`#Selection of the features of interest part of the learning zone building`
`?foi a ssn:FeatureOfInterest; dul:isPartOf <http://www.semanticweb.org/corentin/ontologies/2016/11/urban_sustainability_assessment#ebbwvale/learningzone>.`
 	
`}`
	:ebbwvale/learningzone_heat_energydemand
	
	:ebbwvale/learningzone_air_indoortemperature
	
	:ebbwvale/learningzone_air_radiantemperature
	
	:ebbwvale/learningzone_air_indoorrelativehumidity
	
	:ebbwvale/learningzone_electricity_energydemand
	...

In Query 10, the FILTER construct is used in order to keep the features of interest that related to "Heat" within the Learning Zone and their related observation.

Query 10. Retrieve the observations related to the learning zone features of interest with "Heat" in their label.

SPARQL query	Output (2016 records in 0.247 sec)

`SELECT ?ob ?foi ?property WHERE{`
 	
`#Selection of the features of interest part of the learning zone building`
`?foi dul:isPartOf <http://www.semanticweb.org/corentin/ontologies/2016/11/urban_sustainability_assessment#ebbwvale/learningzone>;`
`rdfs:label ?foi_label.`
 	
`#Selection of the observation associated to those features of interest`
`?ob a ssn:Observation ; ssn:featureOfInterest ?foi;`
`ssn:observedProperty ?property.`
 	
`#Filter the features of interest with the term "Heat" in their label`
`FILTER regex(?foi_label, "Heat")`
 	
`}`
	:observation/26209/
	:ebbwvale/learningzone_heat_energydemand
	:energyproperty
	
	:observation/26210/
	:ebbwvale/learningzone_heat_energydemand
	:energyproperty
	
	:observation/26211/
	:ebbwvale/learningzone_heat_energydemand
	:energyproperty
	
	:observation/26212/
	:ebbwvale/learningzone_heat_energydemand
	:energyproperty
	...

Then in Query 11, the observation values associated with those features of interest are queried. This showcases how objects present in a BIM or cityGML model can be linked to observation values stored in the RDBMS.

Query 11. **Retrieve values, times and unit of the :ebbwvale/learningzone_heat_energydemand**

SPARQL query

```
SELECT ?ob ?value ?time ?unit WHERE{

#Selection of the features of interest part of
the learning zone building
?foi dul:isPartOf
<http://www.semanticweb.org/corentin/ontologie
s/2016/11/urban_sustainability_assessment#ebbw
vale/learningzone>; rdfs:label ?foi_label.

#Selection of the observation associated to
those features of interest
?ob a ssn:Observation ; ssn:featureOfInterest
?foi; ssn:observedProperty ?property.

#Filter the features of interest and their
observations with the term "Heat" in their
label
FILTER regex(?foi_label, "Heat")

#Retrieval of the values, times and unit
associated with the filtered observations
?ob ssn:observationResult [ssn:hasValue
[:hasNumericValue ?value];  qudt:unit
[qudt:baseUnitDimensions ?unit]];
ssn:observationResultTime
[dul:hasRegionDataValue ?time].


}ORDER BY ?time
```

Output (2016 records in 0.898 sec)

```
:observation/26209/
"2015-09-
17T00:00:00+01:00"^^xsd:dateTim
e
"18"^^xsd:decimal
"kWh"^^xsd:string

:observation/26210/
"2015-09-
17T00:30:00+01:00"^^xsd:dateTim
e
"18"^^xsd:decimal
"kWh"^^xsd:string

:observation/26211/
"2015-09-
17T01:00:00+01:00"^^xsd:dateTim
e
"18"^^xsd:decimal
"kWh"^^xsd:string

:observation/26212/
"2015-09-
17T01:30:00+01:00"^^xsd:dateTim
e
"18"^^xsd:decimal
"kWh"^^xsd:string

:observation/26213/
"2015-09-
17T02:00:00+01:00"^^xsd:dateTim
e
"18"^^xsd:decimal
"kWh"^^xsd:string
...
```

In Query 12, the KPIs present in Ebbw Vale are investigated. They are seen here as the result of some sensing methods (e.g calculation methods), themselves participating in sensing events that take place in Ebbw Vale. They are filtered using the tag “Indicator” as simple measurements, criteria, subthemes and themes are equally seen as outputs of some sensing methods.

Query 12. **Which KPI are present within Ebbw Vale?**

SPARQL query

```
SELECT DISTINCT ?label WHERE{


#Find the list of the present KPI as an output of a
sensing event in ebbw vale
?sensingevent a :SensingEvent; dul:hasParticipant
?sensing,
<http://www.semanticweb.org/corentin/ontologies/201
6/11/urban_sustainability_assessment#ebbwvale>.
?sensing a ssn:Sensing; ssn:hasOutput [rdfs:label
?label].
```

Output (14 record in 0.261 sec)

```
"AirQualityLevelIndicator"
^^string

"CarbonEmissionsIndicator"
^^string

"ComplaintRegardingThermal
ComfortIndicator"^^string

"CostEffectiveDevelopmentI
ndicator"^^string
```

`FILTER regex(?label, "Indicator").`
 	
`}`
	"ElectricalLossesIndicator"^^string
	
	"TotalEnergyDemandIndicator"^^string
	...

Now aware of the presence of the *TotalEnergyDemandIndicator* from the previous query, Query 13 investigates how the KPI scores. It is considered as the observation results of some observation and associated with a time and value.

Query 13. Retrieve scores and times of the Total Energy Demand Indicator.

SPARQL query	Output (2016 records in 3.938 sec)

`SELECT ?kpi ?time ?value WHERE{`
 	
`#Selection of the Total Energy Demand KPI in the ebbw vale urban system`
`?kpi a :TotalEnergyDemandIndicator;`
`:inUrbanSystem <http://www.semanticweb.org/corentin/ontologies/2016/11/urban_sustainability_assessment#ebbwvale>.`
 	
`#Selection of the observations that have for output this KPI and the time associated each the observation`
`?ob a ssn:Observation; ssn:observationResult ?kpi; ssn:observationResultTime [dul:hasRegionDataValue ?time].`
 	
`#Retrieve the score of the Total Energy Demand KPI`
`?kpi :hasAbsoluteScore [:hasNumericValue ?value].`
`}`
	:sensoroutput/108866/ "2015-09-17T00:00:00+01:00"^^xsd:dateTime "100"^^xsd:decimal
	
	:sensoroutput/108868/ "2015-09-17T00:30:00+01:00"^^xsd:dateTime "100"^^xsd:decimal
	
	:sensoroutput/108870/ "2015-09-17T01:00:00+01:00"^^xsd:dateTime "100"^^xsd:decimal
	
	:sensoroutput/108872/ "2015-09-17T01:30:00+01:00"^^xsd:dateTime "100"^^xsd:decimal
	...

In Query 14, the benchmark used for the *TotalEnergyDemandIndicator* score determination is retrieved, giving a better understanding of the KPI value. In the present case, the benchmark and its value are stored in the RDBMS. However, it is feasible to have the benchmark pointing to a public URI, implemented by a governmental institution, with an up-to-date value. This would allow the system to keep up to date records of reference values for the KPIs calculation.

Query 14. Retrieve the benchmark associated with the Total energy demand indicator.

SPARQL query	Output (1 record in 0.906 sec)

<pre>SELECT DISTINCT ?ben ?benchvalue ?unit WHERE{
 	
#Selection of the Total Energy Demand KPI in the ebbw vale urban system
?kpi a :TotalEnergyDemandIndicator; :inUrbanSystem <http://www.semanticweb.org/corentin/ontologies/2016/11/urban_sustainability_assessment#ebbwvale>.
 	
#Selection of the observations that have for output this KPI and the time associated each the observation
?ob a ssn:Observation; ssn:observationResult ?kpi; dul:satisfies ?goals.
 	
#Retrieve values and unit associated with the benchmark
?ben a usa:Benchmark ; ssn:hasValue [dul:hasRegionDataValue ?benchvalue; qudt:unit [qudt:baseUnitDimensions ?unit]] ; dul:expresses ?goals.
 	
}</pre>
	:benchmark/60/
	900
	"kWh"^^string

Query 15 examines the features of interest and their properties at the origin of the *TotalEnergyDemandIndicator* calculation. Indeed, if in Query 12, the *TotalEnergyDemandIndicator* is seen as the result of a sensing method, here the inputs of this same sensing method are investigated. The result is other sensing methods themselves associated with a feature of interest and a property via the observation.

Query 15. What are the features of interest and properties involved in the calculation of the Total Energy Demand Indicator?

SPARQL query	Output (10 record in 0.466 sec)

<pre>SELECT DISTINCT ?in ?foi ?prop WHERE{
 	
#Find the inputs involved in the calculation of sensoroutput/108866/ (a TotalEnergyDemandIndicator)
?sensing a ssn:Sensing; ssn:hasOutput [dul:isExpressedBy <http://www.semanticweb.org/corentin/ontologies/2016/11/urban_sustainability_assessment#sensoroutput/108866/>]; ssn:hasInput ?in.
 	
#Find the observation associated to those inputs and their related feature of interest and property
?obs ssn:sensingMethodUsed ?in;</pre>
	:sensing/1/
	ebbwvale/energycenter_electricity_energydemand
	Energyproperty
	
	:sensing/2/
	ebbwvale/energycenter_heat_energydemand
	Energyproperty
	
	:sensing/8/
	ebbwvale/generaloffice_electricity_energydemand
	Energyproperty
	
	:sensing/9/
	ebbwvale/generaloffice_heat_energydeman
	Energyproperty

ssn:observedProperty ?prop; ssn:featureOfInterest ?foi. }	:sensing/13/ ebbwvale/learningzone_electricity_energydemand Energyproperty …

Finally, following Query 15, Query 16 retrieves the buildings relating to the features of interest used for the *TotalEnergyDemandIndicator* calculation. In this way, one can know which real object are directly or indirectly involved in each KPI determination.

Query 16. **What are the buildings involved in the calculation of the Total Energy Demand Indicator?**

SPARQL query	Output (5 record in 1.607 sec)

`SELECT DISTINCT ?build WHERE{`
 	
`#Find the inputs involved in the calculation of sensoroutput/108866/ (a TotalEnergyDemand Indicator)`
`?sensing a ssn:Sensing; ssn:hasOutput [dul:isExpressedBy <http://www.semanticweb.org/corentin/ontologies/2016/11/urban_sustainability_assessment#sensoroutput/108866/>]; ssn:hasInput ?in.`
 	
`#Find the observation associated to those inputs and their related features of interest and properties`
`?obs ssn:sensingMethodUsed ?in; ssn:observedProperty ?prop; ssn:featureOfInterest ?foi.`
 	
`#find the building relating to those features of interest`
`?build a :Building.`
`?foi a ssn:FeatureOfInterest; dul:isPartOf ?build.`
 	
`}`
	:ebbwvale/leisurecenter
	
	:ebbwvale/learningzone
	
	:ebbwvale/energycenter
	
	:ebbwvale/school
	:ebbwvale/generaloffice

Query 17 gives an example of the CONSTRUCT structure in SPARQL, allowing the creation and storage of generated triples from a SPARQL query. In the present example, the administrator of the RDBMS is allocated a location in Ebbw Vale General Offices, information that was not present neither in the triple store nor the RDBMS.

Query 17. **Assertion of the location of the MySQL admin in Ebbw Vale General Office.**

SPARQL query	Output (5 record in 1.607 sec)

`CONSTRUCT {`
`?admin dul:hasLocation http://www.semanticweb.org/corentin/ontologies/2016/11/urban_sustainability_assessment#ebbwvale/generaloffice`
`}`
 	
`WHERE{`
 	
`?admin a dul:Person; dul:hasDataValue "kusterc@cardiff.ac.uk".`
 	
`}`
	# Object property assertion axioms
	:person/1/ hasLocation generaloffice .

6.7 USA FRAMEWORK WEB APP ARCHITECTURE

6.7.1 Architecture

The USA framework has been designed with full integration within a web application in mind. That would give its users seamless access to the sensor data and discovery of the key performance indicators. The system is composed of 3 layers with the data layer, application layer and user interface layer, as illustrated in Figure 6-18. The data layer includes 4 different databases servers:

- an RDBMS such as MySQL or PostgreSQL where sensors' data and metadata are stored;
- an RDF store such as Apache Jena or GraphDB where RDF instances of BIM models and cityGML models can be found as well as benchmarks fetch from governmental open RDF repositories and any other relevant information;
- A BIM server such as ifcBIMServer or openBIMServer where BIM models are stored;
- A Geo database such as 3DcityDB where cityGML models are stored.

Some pieces of information will be redundant across all the servers (e.g. a sensing device). Therefore, the connection is implicitly made between them via the allocation of a unique ID across all the platforms.

The application layer includes:

- the RDBMS client for java Hibernate, an object-relational mapping tool that enables the manipulation of databases in Java. This feature is optional and was meant for administrators that wish to re-use the native RDBMS schema presented in Figure 6-3 and to dump their sensor data in an easy manner. One can also use custom RDBMS.
- the OBDA with an implementation of ONTOP that connect to the RDBMS via JDBC protocol and to the RDF store by loading the ontology in ".owl" format. If one re-uses the native RDBMS, the OBDA mappings developed in section 6.5.1.2 can be re-used as-is. If not, he or she will have to design new OBDA mappings between its own RDBMS and the USA ontology.
- a BIM client that allows the system to parse BIM entities IDs from the 3D BIM interface to a SPARQL query. In this way, the user can use the 3D BIM interface to discover BIM entities and generate SPARQL queries.

– a cityGML/GIS client that, similarly to the BIM client, allows the system to parse cityGML entities IDs from the 3D city interface to a SPARQL query.

A user interface layer that includes:

– a sensor data entry user interface that allows administrators to store sensor data and meta-data into the native MySQL DB easily.
– a SPARQL endpoint that allows the user to query the system and discover sensor data and KPI. On top of which several libraries can be implemented in order to visualise the query output into graph, maps, trees, tables etc.
– a 3D BIM user interface that allows to navigate into 3D BIM models and discover building objects IDs that can subsequently be used into the SPARQL queries.
– a 3D city model user interfaces that in the same fashion allows to navigate 3D cityGML models and discovers city objects IDs.

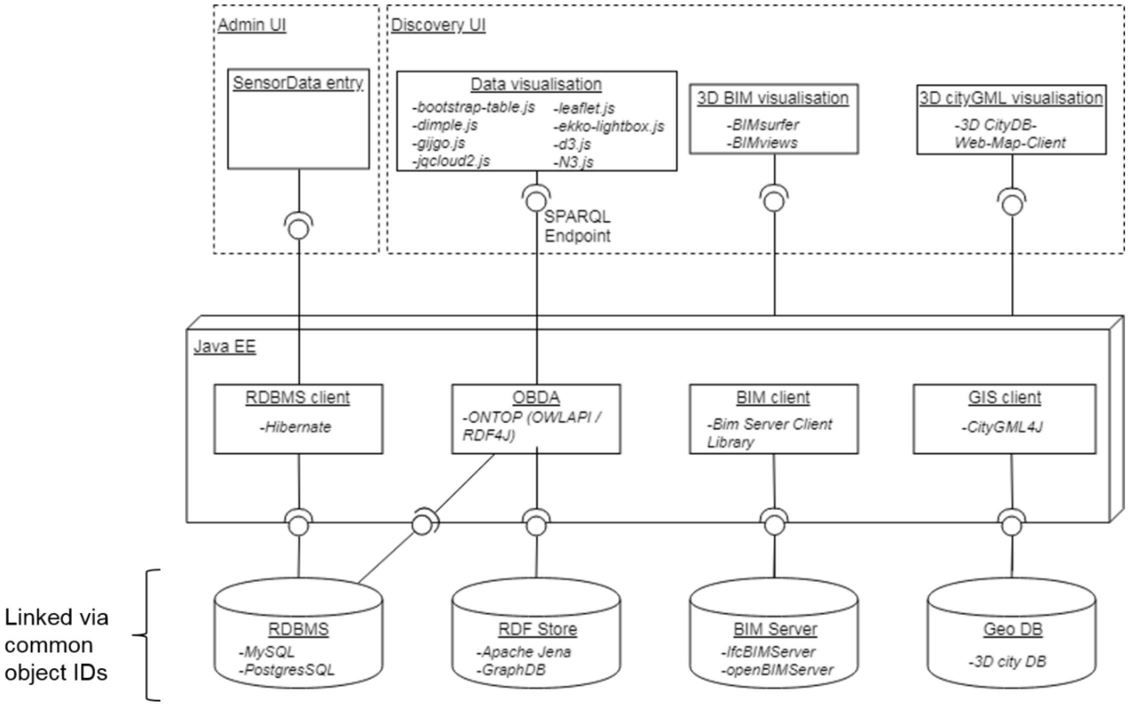


Figure 6-18 framework web app architecture

Overall, this architecture is believed to enable publishing sensor data as well as navigating urban systems 3D models in order to ease the design of SPARQL queries that retrieve the most relevant information. The user can develop sophisticated queries and infer information in real-time in an easy manner.

6.7.2 End User Interface

The USA platform would greatly benefit from the integration of a user-friendly interface that allows users that are not familiar with SPARQL queries to handle the tool in the most efficient way. In the prospect of an inclusive tool, there is a clear need to provide information in the most transparent and user-friendly way. Dashboards are nowadays widely used for KPIs' visualization in Business Information [654]. In the energy sector, the increasing use of ICTs-based tools led to a higher amount of data and therefore to question the good use of dashboards [655]. The development of dashboards must be in line with the vision and goals of the user [654]–[657]. Yun et al. developed a dashboard in order to raise awareness and encourage office workers towards more environmentally sustainable behaviour [658]. Based on their researches, they designed a dashboard with aggregated real-time and historical data by using bars, areas and line charts, they provide short and long term recommendations, let the possibility to control items from the dashboard and finally provide a report of the progress and the saving realized. Those features are mentioned as “Self-monitoring and Comparison”, “Recommendations”, “Control” and “Reward” and are considered as essential to engage people. Dashboards are also used in the context of decision making [655]. This configuration requires the integration of features enabling the visualization of strategies such as providing goals, alerts, analysis report, enabling only the relevant visualizations for a specific role, and sorting them according to priorities and semantic criteria [654].

Kumaraswamy et al. recommend the use of widgets as visual aids. They integrate media such as pictures and videos, and form elements such as drop down menus, check boxes and date pickers in order to take user inputs [656]. Kintz advocates for the use of bar chart for categorized data, line chart for comparison over the time or distribution, number or sparkline for a single value, and bullet graph for actual value vs objective [654].

Dashboards should equally take in consideration a possible use on mobile computer (smartphone, tablet) [659] and the visualization of inter-groups interactions for clustered information [660].

Overall, the identification of users' focus, the aggregation of information, the use of charts, tables, gauges, the comparison with references point, the provision of recommendation, alerts, reports, consideration of user input are features supported by many studies toward KPIs' dashboards design.

6.7.2.1 *User interfaces requirements*

Before proceeding to the design of the user interface, one must consider the set of requirements and goals of the USA framework. This set of requirements is in line with the USA Ontology Requirement Specification presented in Section 5.2, especially the definition of the goals, domains, scopes, uses and users. The tool provides access to the sensor measurement as well as the KPIs calculated. One must be able to explore a certain urban system or neighbourhood and access building, appliances and sensor characteristics easily. Performances must be comparable in order to help decision making. Additionally, linkages between physical elements and indicators would help in this task.

Users can be divided into two categories: the administrators that are in charge of publishing the sensors data and meta-data within the USA framework and the explorers that are the direct user of the application for data discovery. Consequently, in term of design, two distinct interfaces can be investigated: an administrator interface with elements that allow an easy input of sensor data and meta-data; an explorer interface where the user can proceed to queries, filter the information, visualise it via dashboards and produce reports in different formats.

6.7.2.2 *The administrator interface*

If a user wants to build the software from scratch using its own RDBMS, one can do so. However, this will requires re-designing the ONTOP mappings so that they fit the integrated database. Same goes for the KPI calculator module that has been designed to fetch sensor data over USADB.

Therefore an administrator interface could ease the publishing of sensors data by a user after he or she deploys the framework on his/her server(s). Figure 6-19 shows a possible design for the administrator interfaces. On the first page, the administrator will have to enter its credentials to enable the publication of sensors data. Once logged in, the administrator can scroll down the list of available sensors, fetched from the BIM or cityGML servers and can associate metadata such as the feature of interest, property and unit. Information about relative and absolute position can be taken directly from the models. Then, one must enter the sensor connection settings so that the live stream of data can be dump into the MySQL database in near real-time.

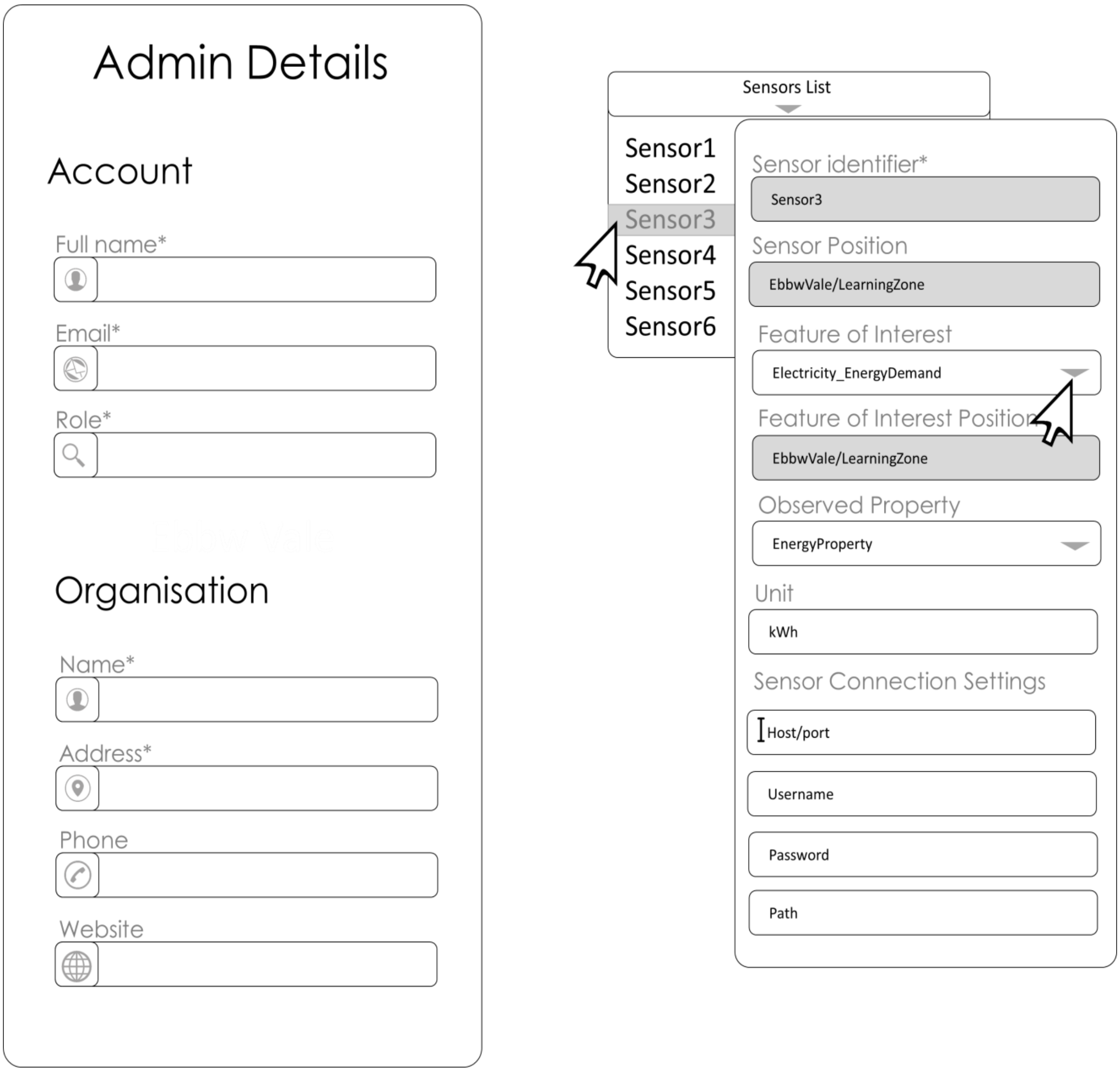


Figure 6-19 Administrator interfaces

An alternative to this administrator interface would be the creation of custom properties allocated to sensors objects into the BIM and/or cityGML models that could be parsed, aligned with the Hibernate objects in JAVA and used subsequent storage in the RDBMS.

6.7.2.3 *The explorer interface*

The explorer interface has been designed based on the Wikidata Query Service (WDQS) [661], a tool that provides a SPARQL API for users to query the Wikidata data. Figure 6-20 shows how the 3D models can be explored using 3D visualisation schemes such as BIMsurfer, BIMviews, 3D cityDB Web Map client. This 3D interactive model is useful to extract object IDs and help to design SPARQL queries. For instance, in the following example, a pop-up window shows all the sensors and their IDs present in the Learning Zone building. When selecting a specific sensor, its ID is printed into the SPARQL query box so that the user can build query related to it. In the present example, the user retrieves the values measured by the Sensor14 corresponding to the electricity demand of the building. The result is presented in a tabular form and can be exported into various formats such as CSV, JSON etc.

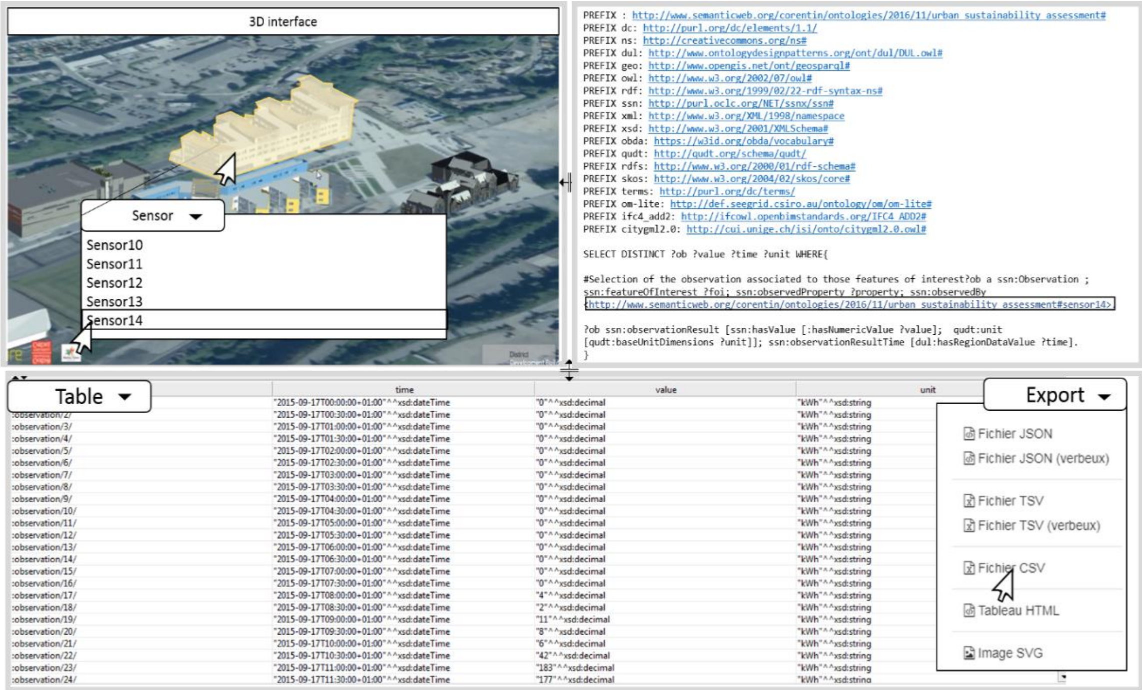


Figure 6-20 User interface view 1

Figure 6-21 gives a second view and showcases how the result can be displayed in a different form, for instance as a line graph. This feature is inspired by the WDQS where numerous results views are efficiently integrated within the user interface [662]. In the service, the result formats are analysed so that the system displays the result in the most appropriate form, coordinates will most likely be displayed on a map, time series on a line chart etc.

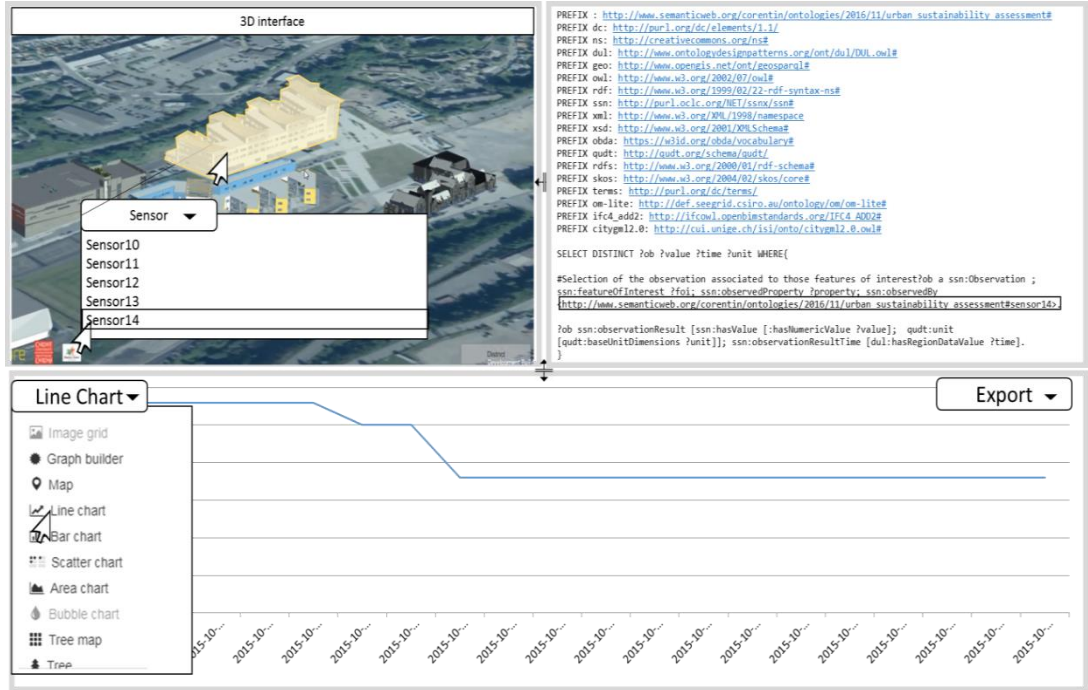


Figure 6-21 User interface view 2

Finally, in this last query, the user investigate the different KPIs calculated within the Ebbw Vale urban system and display them into a grid of doughnut charts with their scores and label.

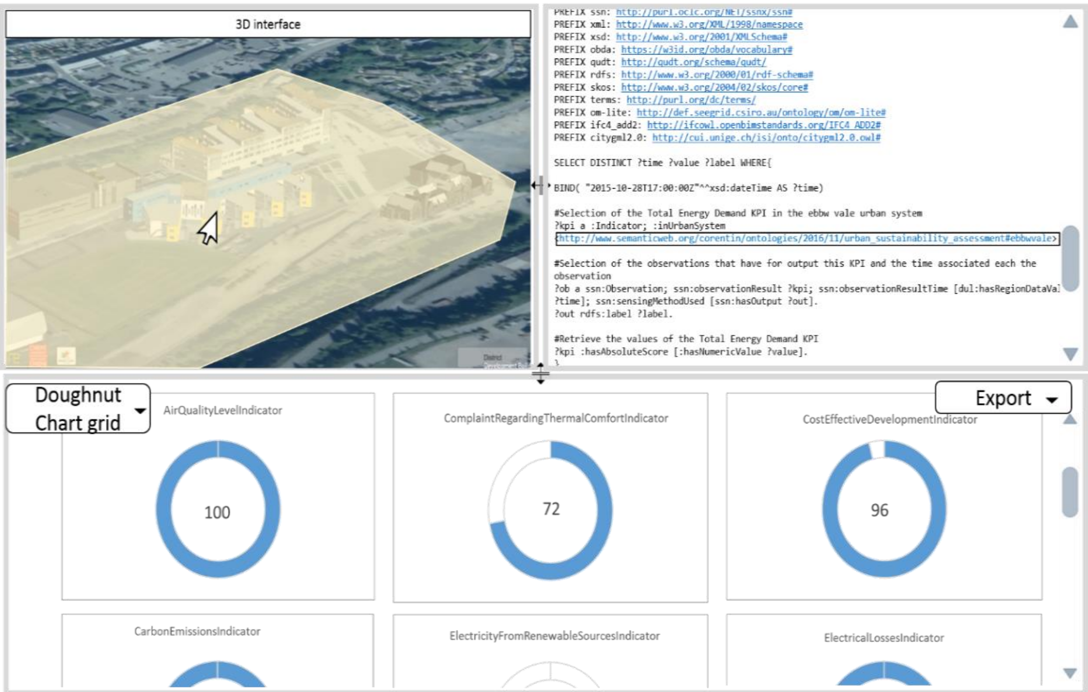


Figure 6-22 User interface view 3

Overall, such user interface is believed to help in the construction of SPARQL query and the discovery of sensor data as well as the easy visualisation of key performance indicators.

Additional features could be integrated such as the "query helper" also found in the WDQS that allows inexperienced users to develop SPARQL queries via templates [663].

6.8 SUMMARY

This chapter demonstrates the validity of a real-time urban sustainability assessment approach via the implementation of an Ontology-Based Data Access system for the discovery of a selected few KPIs. The case study of the Ebbw Vale neighbourhood, the Work, has been used for this demonstration. The OBDA has been developed via the ONTOP application that allows its user to create mappings between a RDBMS and an RDF store. The Java-based software also provides a SPARQL end-point from which a set of testing queries have been implemented. Despite certain limitations due to the nature of OBDA and ONTOP, the system shows good performance and proves the possibility of solving heterogeneous sources linking BIM data and sensors data and to develop complex queries. Finally, the web application architecture and front-end design have been addressed, giving guidelines for the future development of the application.

Overall, the chapter has provided evidences to answer the following research questions:

RQ2. How an effective urban sustainability assessment can help different parties of a city in their decision making?

With the help of the case studies, this chapter has demonstrated the possibility to answer complex queries over the USA frameworks leveraging on the OBDA system. It has showcased the usefulness of a real-time urban sustainability assessment approach where one could efficiently discover KPIs, retrieve data in real-time, forecast future trends and link information to the urban environment abstraction present in BIM or cityGML models. All those feature will help in the decision making at the urban level.

RQ3. How can sustainability assessment leverage the smart city paradigm, specifically ICTs and the IoT?

In this chapter, the author has showcased the feasibility of KPIs determination in real-time based on sensor networks deployed in several case studies. It described in details the data flow from the sensor acquisition to the user interface, explicating how raw data could be used to define sustainability KPIs at the urban level.

RQ4. How can semantic web technologies unify heterogeneous data resources for holistic services and applications?

Queries have been testing the linkage between BIM or cityGML model and real world sensor data from the different case studies. Moreover, federated sparql queries have been discussed, giving the possibility to retrieve information from a singular sparql endpoint. The developed web service has demonstrated the usefulness of semantic web technologies to deal with heterogeneous information system by setting a common semantic understanding over the data produced.

RQ4. How technological, human and financial assets relate to such service provision approach in the smart city paradigm?

The choice of an open web service with a user- friendly interface advocates in favour of an accessible, transparent and participatory design. In a broader perspective, it reflects an actual need in the smart city movement of citizen centred services with an open and collaborative dimension.

Discussion and Conclusion

In this section, the author will open discussion on the different outcomes of the study. After summarizing the different approaches taken and the research findings, each of them will be discussed and the relevant contribution to the body of knowledge will be highlighted. Therefore, the current USA schemes analysis, the new USA scheme created, the investigation of assets for KPIs real-time measurement, the semantic web technologies to deal with data sources interoperability and the final USA application are presented. Lessons from projects and interrogation on the industry reception of such solution will be discussed as well. Then, the contribution to knowledge will be clearly outlined by answering the research questions formulated in the Introduction. In a third section, the author will present the limits of the current solution and open on future efforts that need to be carried on. Finally, the author will conclude on a closing remark on the domains studied and the contribution of this study.

7.1 DISCUSSION OF THE RESULTS

7.1.1 USA schemes findings

The literature review has highlighted a number of limitations of the current urban sustainability assessment schemes. No less than 29 frameworks [16], [17], [59]–[68], [50], [69]–[77], [51]–[55], [57], [58] have been analysed and compared in details, looking at their individual structure, indicators, criteria and themes addressed, their weighing system and their overall applications. Those frameworks have been designed by experts in their own field. There is no doubt that the current frameworks addressed relevant indicators that are the product of research and reflection on the field of sustainability at the urban level. Most of these frameworks have already been applied by the construction industry to evaluate newly developed projects across the world. They are supported by governmental institutions which legitimate their value of referenced frameworks. However, some criticisms have emerged in the academic community towards the biases that some of the frameworks apply [24], [25], [46], [47]. Indeed, it seems that most of these frameworks have been designed by and for the construction industry. In that regard, it seems that the frameworks tend to focus on aspects that are the main concerns of the industry such as material logistic, energy savings, environmentally efficient construction etc. As such, the study of the different frameworks has shown that environmental concerns were most likely to be addressed than socio-economic issues for instance. Additionally, it seems that most of those frameworks have been designed to be applied during the design stage of development, the result of an industry-oriented perspective. There are relatively few schemes that cover sustainability assessments during the operative stage of a neighbourhood or urban area. Those observations go along some concerns on the motivation to use those frameworks to simply "*seek market appeal on projects*" [24]. Consequently, there is room for a new type of assessment framework to be applied during the operative stage of development that would consider all the parties involved in a greater extent, including district managers, inhabitants, experts etc. Such frameworks (applied during the operative stage) would benefit from a more user-centric approach, redefining the target user with greater participation of all the stakeholders and not only the construction and developer professionals.

Hopefully, the participation of heterogeneous bodies will help to converge toward a definition of what does sustainability at the urban level, reducing biases and reaching a greater consensus that lacks at the moment.

It could equally help to solve another issue highlighted in the review of the current schemes which is the problem of adaptability. Following the region where the assessment is applied, certain aspects need to be stressed more than some others as they might be the source of greater concerns. In the existing instances, local adaptability is either absent, which mean the assessment has been designed for one particular location; or it is tackled in an ad-hoc manner where, after an extended study, the experts decide which aspects to stress. If a personalized adaptation is often the most relevant way to conduct the assessment in different locations, it could greatly benefit from automatized mechanisms to help decide on the elements to pressure. Therefore, integrating more parties into the loop could help in gathering further feedbacks and information which ultimately help to adjust the assessment to a particular scene.

In addition to local adaptability, Orova mentions changes over time [47]. This aspect is particularly essential when considering an evaluation during the operative stage. Indeed, a neighbourhood is a complex system and a fast-changing environment where certain aspects will require a continuous evaluation in order to meet strict requirements that ensure sustainability. Examples such as energy, water or waste management must consider production and demand in near real-time in order to deliver the most efficient, economic and environmentally-friendly service. Mostly intended for new urban development design, the current USA frameworks lack dynamism. Consequently, the perspective of developing a new USA framework that would cover urban areas in operation has led to the conceptual re-definition of the assessment, especially matters on the temporality in which it fits. Section 2.1.3.2 of the literature review developed this notion by investigating background on buildings (and neighbourhoods by extrapolation) good operation and maintenance, opening on the idea that recurrent evaluation and efforts would lead to greater performance and lifespan. Moreover, in a context where governmental leaderships release increasingly strict regulations, it is essential for assessment schemes to be able to adapt to the newly introduced requirements.

Considering those aspects, the research operates a conceptual shift from "classical" assessment to a decision support system providing its users with the best solutions to operate and maintain an urban area. This idea is supported by studies such as Coplak and Raksanyi [420] or Sharifi [24] that states that a sustainability assessment has vocation to help to take the most appropriate decisions and not simply awarding labels. Consequently, decision support systems have been further investigated in Section 2.5, highlighting some key aspects, in particular the provision of live data and efficient forecasting models.

Forecasting models have been the subject of a systematic literature review in order to identify the most commonly used and how they have been set. Taken in the context of decision support for electrical load management, the author has evaluated their relevance regarding the input variables used, the scale, time resolution and time horizon. The main output is a taxonomy presented in section 2.5.3.3 that recommends the use of Artificial Neural Network, Support Vector Machine or Time Series Analysis (ARMA, ARIMA, SARIMA …) for very short to short term prediction (1min to 1 week) while regressions are often sufficient for long term prediction and low log frequency. It is equally advised to pre-process the inputs data by filling the gaps, smooth the trends and analysing their correlation with the output. Such information is essential in the pursuit of the development of a decision support system.

Live and historical data are at the core of decision systems, being the source of knowledge indispensable to discover insights on the observed object. It is there that lies the importance of the smart city paradigm and of what the IoT has to offer.

7.1.2 The New USA Scheme

In the research of a decision support system that would assess in real-time sustainability, a new scheme has been developed. The development of such scheme has required the analysis of the already existing schemes. The author has gathered the most encountered indicators and created a new scheme as the synthesis of those. By selecting the aspects where the existing schemes converge, the objective was to build a basis framework that does consensus within the expert community. A top-down approach has been employed looking first at the main themes addressed then subthemes and criteria. Those have been categorised and sorted resulting in the creation of a mind map with different levels of abstraction. This mind map helped in guiding the indicator selection on more specific aspects. In the end, 193 indicators have been selected, fitting into 90 criteria, 25 subthemes and 8 themes.

Beyond the simple selection of indicators, the feasibility of measuring and employing them in real-time has been investigated. A DELPHI consultation has been initiated to validate and gather experts' opinions on the validity and value of such approach. More than 250 experts were approached to answer the DELPHI consultation within 2 pilot versions and a final version with 2 rounds of questions; for a participation rate between 10% and 15% following the versions and stages. The low participation rate has constrained to downscale the original framework to a few selected indicators. The indicators' selection for validation has been done on the basis of the data available in the case studies which would guarantee their use in a subsequent experimental validation. Overall, on the 193 KPIs originally present, 17 have been selected, namely Air Quality Level / Complaints, Carbon Emissions, Complaint Regarding Thermal Comfort, Cost Effective Development, Electrical Losses, Electricity From Renewable Sources, GHG Emissions, Heating From Renewable Sources, Heat Losses, Onsite Energy Generation, Organic Waste For Energy Generation, Thermal Gradient Differences, Time out of Thermal Comfort Levels and Total Energy Demand. The first survey round was focusing on 3 main aspects that are the overall relevance of the indicators in sustainability assessment, the ability to capture them in real-time and the degree on which they can be improved through managerial and operational efforts. The responses showed that at least 66% of all the experts involved considered the KPIs as relevant while the majority of them think they are measurable in real-time but it can be costly and that they are possibly improvable via operational and managerial efforts. In a second round, the log frequency and forecast time horizon of each of the 17 KPIs have been investigated. Most of the experts think that hourly to daily log frequencies are enough to consider the measurement as “in real-time”, at the exception of *OrganicWasteForEnergyGeneration* with suggestions of weekly to monthly measurements. About forecasts, the experts estimate that at least daily forecasts are required or later for *costs*, *EquippedBuildings* and *OrganicWasteForEnergyGeneration.*

In addition, the study of feasibility has led to the review of the literature to seek actual instances of means of measurement for the newly developed framework. Sensors, remote sensing, GIS, BIM, open government platforms or crowdsourcing and data mining are as many technologies that could support the determination of the KPIs in real-time. They are all part of a bigger picture that includes the IoT, ICTs and the smart city paradigm. These aspects are developed in the following section.

7.1.3 Smart City Assets for sustainability assessment

7.1.3.1 *Sensors*

In 2018, half of the 20 billion connected devices (including more conventional ones such as fixed and mobile phones, computer and tablets) sold were IoT devices [123]. A figure on the rise and by 2022 IoT will outnumber mobile phones, computers and tablets. With the proliferation of connected devices, goes along their increasing capability to capture real-world phenomenon. Examples are numerous with the appearance in recent years of smart clothing and wearable devices to measure physical activities and health [138], [139], smart transports devices that record traffic and passenger frequencies [539]–[541] or smart home devices that can record the indoor air quality [556]. Those new technologies add up with already existing and well-installed ones capturing energy, water or waste consumption, noise levels, fire alarms, audio recording, video cameras etc. A large number of real-world physical phenomenon can now be monitored in real-time with high precision. Section 0 has shown that out of the 193 indicators, 63 can be measured via sensing devices including, indicators such as total energy demand, electrical losses, water leaks, waste recycled, sky glow, air quality, vehicle miles travelled, fire and crimes records, ICT efficiency, species diversity etc. Those are supported by 24 references that showcase the feasibility of such measurements in the case of the most “challenging” indicators.

7.1.3.2 *GIS and remote sensing*

In addition to “direct” sensing, the research has highlighted the importance of remote sensing coupled with up-to-date GIS technologies and GPS able to capture large scale phenomenon and social constructs. Indeed, in the last decades, the field of remote sensing has dramatically improved with the sharp enhancement of aerial and satellite imagery that has reached a precision within the centimetres scale [191]. Other technologies such as LiDAR have participated in the production of accurate systems able to detect certain object for instance [196]. Objects detection and identification studies have increased, leveraging on high-resolution images used to train machine learning algorithms [192]–[195]. Those technologies combined can help in monitoring certain KPIs addressed in the USA scheme. They have been found to be useful for 65 out of 193 indicators with the ability to measure solar panels energy potentials, GHG emissions, spaces functions and characteristics (open, green, recreational, brownfields etc), transport penetration, distances to services and facilities or detection of urban furniture such as bicycles racks, parking lots, crosswalks etc. Twenty-two publications have been taken to illustrate such application of the technology. There is a clear potential for the development of 4D GIS able to capture temporal changes with frequent data acquisition [211], [212]. Opening such system and enabling some sort of participatory approach could allow crowdsourcing of information that would ensure an increased level of accuracy.

7.1.3.3 *BIM*

The AEC industry is considering with a greater extent the use of BIM models for their development project. They have acquired a central role in the construction supply chain enabling the management of assets and processes [254], [255]. Originally used as a design tool, 3D BIM models include more and more features, listing a complete inventory of the multiple building objects, allowing clash detection, multi-disciplinary coordination and a user-friendly 3D rendering. The field is moving toward “nD BIM”, extending its capabilities by integrating schedules and progress monitoring, cost and sustainability analysis, facility management or lifecycle assessment [664]. In the context of building in operation, as-is BIM models are essential to guarantee the relevance of the facility management and the integrity of its components [163]. The literature review has provided lead in the pursuit of as-is BIM production with the integration of technologies such as RFID and barcoding of building components and appliances [151], [175], [176], 3D scanning of building geometries and equipment [177]–[180] or the use of high resolution imagery and photogrammetry [181]–[183] [184], [185]. Such technology is believed to help in the monitoring of 17 KPIs including features such as electricity reliability, the provision of certain appliances such as composting unit, materials used, development costs, spaces areas, shadings, building recovery, building envelope efficiency and others.

7.1.3.4 *Open government databases*

The assessment of certain socio-economic aspects requires the collection of sensitive information about citizens and organisations that none of the technologies previously cited can catch. The sensitive and private nature of information on revenue, health, consumption or movement requires a trustworthy party that the government embodies [407]. Without any profit interests, governments have the privileged position to ensure the collection of such information. This trust must be maintained and strengthened via complete transparency of the information usage, the sovereignty of the citizens over data and the provision of open platforms to access the information [402]. Moreover, governmental institutions are often in charge of services provision on waste collection, transport, education, care programs etc, which re-enforce its position to collect and distribute such type of information. Consequently, a certain number of sustainability indicators could benefit from the implementation of open governmental data platform and APIs. There are in the new USA framework, 45 KPIs that can be determined in this way including information on waste collection, vehicle emissions, education enrolment, income distribution, employment rates, women/men parity in institutions, transport affordability, health care protection etc.

7.1.3.5 *Crowdsourcing and data mining*

Crowdsourcing and data mining are the latest forms of data resources that can be used for sustainability assessment. Those technologies leverage on online platforms and social media to gather information from and on people. The power of crowdsourcing lies in the "wisdom of the crowd" [233] and the production of a high volume and quality knowledge base [236]. Whether it is through incentive or not, citizens are at the source of the information which allows them to actively participate in the evaluation. Data mining can be seen as a passive form of crowdsourcing where online resources (e.g. social media) are scanned and analysed in order to retrieve trends and extra pieces of information from the data mass [224]. Twenty-one indicators could be determined using crowdsourcing and data mining with for instance KPI on plants and species inventory, satisfaction over public services, measuring safety feelings, complaint about air quality, physical activities etc.

7.1.4 Reaching interoperability with semantic web technologies

If all the data sources presented earlier constitute a great support for the assessment of sustainability, they are, however, dispersed and heterogeneous. This is a challenge to overcome for the creation of a holistic system. Smart cities are facing a great impulse but many estimate that they can only meet their true potential if a ubiquitous and holistic system is designed [27], [638], [665], [666]. For that, the technology needs to show interoperability between heterogeneous devices and data. The delivery of standards on data and meta-data format is an option to improve homogeneity and ensure M2M communication. However, the field is fast growing and complex which makes the traditional application of standards difficult [247]. Consequently, a good approach would be the development of technologies to translate and/or understand data formats across systems.

In that context, semantic web technologies can contribute to the creation of an interoperable system. Semantic web technologies have been included in the WWW development agenda since a very early stage. They are at the origin of the creation of linked data and the web 3.0 where the data and information are linked regardless of their formats or specificities [274]. The link is defined via the integration of semantic annotations that give meaning to pieces of information and how they relate one to another. It can be done via the development of ontologies that provide a representational vocabulary for a domain via the definition of classes, relations, functions or other objects [295]. The domain is described with a set of statements formed by triples subject-predicate-object. Those triples are the building blocks of the ontology and allow reasoning engines to infer extra pieces of knowledge. Ultimately, this representational vocabulary can be instantiated with real-world entities providing meaning and intelligent reasoning over the information. The technology is built over unique HTTP URIs that identify each piece of information and from the RDF, RDFS and OWL languages, the languages associated with data, taxonomy and ontology description [300].

In regards to the semantic web technologies capabilities, an Urban Sustainability Assessment ontology has been designed. The ontology development has required the use of a strict methodological approach which ensures the consistency of the information system. Indeed, in order to form a global system of information, the field advocates for the re-use of existing ontologies over certain domains. Therefore, it can be challenging to design an ontology that does not overlay over other existing resources and to find the exact spot where the ontology must seat. Some methodological frameworks exist to help ontologies' designers in defining precisely the boundaries, functions, uses, users and resources to be considered [373]–[375], [377]–[381], [667], [668]. In this research, the NeON methodology has been followed to design the USA ontology. The methodology compiles possible scenarios that the engineer can encounter following the source materials and provides guidelines for each of them. An important step is the ontology requirements specification where the domain, uses and users have been framed. The definition of competency questions helped in this task. The USA ontology has for objective to describe the USA scheme and its related domains, including urban environment characteristics. The ontology is used to discover sustainability KPIs and to better understand how they relate to urban environment components. The user must be able to retrieve values, features of observations, benchmark value, relationships between KPIs and urban objects. Users of the ontology would most likely be district managers and officials in charge of a neighbourhood's maintenance and operation. However, such framework could also be used by inhabitants for a simple KPI discovery and raise their awareness on their neighbourhood performance. Regarding the requirements, a set of ontological resources have been investigated and a selected few integrated within the USA ontology. The ontology is constructed around 4 main modules:

- The observation module that uses the SSN ontology [322] for the description of sensors and their observations
- The KPI module with the description of the different indicators, criteria, subthemes and themes of the USA scheme and how they relate to each other.
- The urban objects module that uses cityGML [357] and ifcOWL [335] ontologies to define city and building level objects that compose an urban area.
- The spatio-temporal module with the use of GeoSPARQL [353] to describe geographical geometries via WKT or GML literals.

In addition, the DUL ontology [324], [325], an upper-level ontology, underlay all the module for additional descriptions of fundamental functions.

Overall, the USA ontology accounts for 884 classes, 113 object properties, 23 data properties and 3751 axioms.

7.1.5 The USA web application

Once the USA ontology developed, it has been integrated as the main information system of a web application that aims to retrieve KPIs and knowledge toward urban sustainability. This application serves as a testing procedure and showcases the feasibility of an ontology embedded system to support urban sustainability assessment. For that, this study has taken the approach of the Ontology-Based Data Access system (OBDA) where a semantic repository overlays a relational database system. In such configuration, the time series produced by the different sensors are stored into the RBD while all other "static" information (e.g. building objects, city objects, agents, units …) are stored in an RDF store. Mappings exist between the two systems to match equivalent entities. Finally, a SPARQL end-point allows a user to query the system and discover insightful information on an urban area sustainability performance.

The development of the USA application has followed the design research methodology where requirements, the product and the expertise improved throughout the iteration. It leveraged on the Ebbw Vale case study which provided extensive district-level data, mostly energy related. Great efforts have been done on the creation of a RDB schema that follows the OGC standards [635] and SSN, the fetching of data from the case study database to the USA database, the selected KPIs calculation and forecasts and the development of relevant mappings.

A set of SPARQL queries has been running to test the performance of the application. The queries were following the competency questions developed during the ontology requirements specification. The application provided answers to the questions on measurement retrieval, released urban objects, KPIs, benchmark, calculation process discovery, geo-location and more in a reasonable time. Complex queries that would have difficulty been answered with traditional information systems, have been answered, showcasing the usefulness of semantic web technologies to bring intelligence and infer information. However, computational time is sensitive to the query design and one must already have a good knowledge of the SPARQL language to come up with performant queries. Same goes for the OBDA mappings that can greatly influence the computation performance. Indeed, if there are several ways to design mappings, one must be careful about creating the optimal design. For instance, some pieces of information can be redundant across several mappings which affect the reasoning task and ultimately the computational performances. Additionally, the application has demonstrated that interoperability between diverse data sources and format is reachable with BIM models information being coupled with sensors data in a unified system.

Finally, some limitation remains especially on the language expressivity and the entailments used by the ONTOP OBDA. Those will furtherly be discussed in section 7.3.

7.1.6 Discussion on Projects Participation

7.1.6.1 *Collaborative Vision*

The projects involved in the action research methodology have provided a great background and insights on the development and deployment of smart solution for cities. All the projects gathered international organisations and experts that aimed collectively to create sophisticated services for energy delivery at the district level. They all highlighted the importance of data in such task and help to picture what smart cities and the IoT can do. The main lesson taken from the projects was the core contribution of a collaborative vision for the success of such services. The different actors must collaborate in order to provide the best possible system. This required a good understanding and agreement on the data ownership and the right over data. This issue is inherent to the deployment of smart services and can be extrapolated from the projects cases to the smart city paradigm as a whole.

In that perspective, some research has been made in recent years on the benefits of "collaborative networks" [279]–[281]. Collaborative networks are defined as the convergence of talents and expertise around a common goal and objectives to deliver a value-added product. Such type of cooperation between heterogeneous organisations is believed to set greater concurrency, workforce performance, domain knowledge dissemination, to reduce environmental impact and trigger innovation [284], [285]. The THERMOSS project has been studied against the collaborative network theory and great evidence showed that the project fits into the collaborative networks domain [669]. Indeed, multiple actors are participating in the operation of service that integrates multiple expertise domains for the completion of environmental, economic and technical objectives. They deployed communication protocols and processes based on existing infrastructures to share information between partners.

Nevertheless, despite the intentions for a greater collaboration, some technical and human issues remain and slow the research and service deployment. An issue highlighted earlier in the thesis is the heterogeneity of the information formalism which requires great efforts of translation and understanding. Indeed, because they have different experiences, work on different software or programming languages, engineers often have tackled the issue in an ad-hoc way to reach interoperability. Solutions must be implemented in the early stages of a project to ensure the homogeneity of the information and help to create a seamless collaboration. Moreover, some technical constraints have been encountered in a project where the data collected by the BEMS were impossible to fetch directly from the sensors. The reason for that was the use by one of the project partner, of a contractor for the sensors deployment that was not part of the consortium. Consequently, even if the project partner had the ownership of the data, it was technically impossible to retrieve the data unless using specific software that allowed collecting a sample of the last few months' measurements. This issue reflects the constraint of data ownership versus technologies' provider. The technologies' provider has the responsibility to provide the data owner means to retrieve his data in an easy manner. Finally, some limitations were due to the lack of trust and scepticism over the solution developed by some of the data and/or infrastructure owners. Some energy district manager did not believe that the performance improvement was worth the time and cost invested in the new solution. Additionally, some inhabitants were really cautious about the use of their personal information. Efforts must be done to clarify how data are being used and the true potential of IoT-based solutions. A complete transparency is important to help the understanding of all the parties involved and bring trust.

One last issue was the competitiveness of certain partners involved the same domain. Indeed, some parties were cautious about sharing information to another party that could be seen as a competitor in their domain. This is an important issue since a city could want to open public-private partnerships with technologies 'providers and/or data owners that are actual competitors. It could affect the actual performance of services provision in smart cities. It is essential to well-defined partners' contributions and to give incentives for collaboration with new business models that define shared risks and revenues.

In that context, it is important to have upper-level organisations that coordinate these common efforts and collaborations to build a greater picture for smart cities and ICTs. In that regard, the input of the European Commission to fund and coordinate the different projects was essential. Indeed, the European Union has a well-defined agenda for innovation and research in promising sectors. It has released detailed roadmaps for Information and Communication Technologies [31] and open data [405]. It helps in defining the parties' contribution and revenue promotes international public-private cooperation via valuable incentives and homogenises efforts over the different domains. Furthermore, the European Union built a regulatory framework such as the General Data Protection Regulation (GDPR) [670] that ensures transparency and ethical use of personal information and brings trust of the data owners.

7.1.6.2 *Semantic web technologies penetration*

An observation from the projects was the limited knowledge and understanding of the semantic web technologies. This issue opens reflections on the current penetration of the technologies within the expert community and more particularly within the private sector. This observation follows the literature review presented in section 2.3.2.4 where experts argue that despite its 20 years' development, the semantic web remains a "niche" and has a hard time to take off as a mainstream standard [310]. Reasons mentioned are the relative immaturity of the domain with inconsistent ontologies developed by small independent groups, the complexity of certain domains and solutions creating a subculture, lack of coordination, the predominant English language, scalability and data protection and credibility. There are still great efforts to be done in the domain for the adoption of such technologies in the private sector. Nevertheless, semantic data modelling for IoT technologies is addressed within the EU Horizon 2020 programme [31] which prefigures an improving situation and push for its greater integration.

7.1.7 Discussion on the Industry Reception

Global warming and the importance of developing sustainable solutions are largely approved within institutions, city developers and the construction industry. The multiplication of sustainability assessment schemes and their application attest it. In parallel, the upsurge of IoT development and smart solutions certifies the smart city paradigm as a seriously considered future by both public and private organisations. More, there are strong evidences showing that sensing devices and data are increasingly valuable and that the data marketplace can be extremely profitable for companies [382], [383]. However, beyond the clear enthusiasm around the domain, there are still some technical and organisational challenges that need to be overcome.

The lessons from the projects open on a bigger discussion on the industry reception on smart IoT-based solutions. The heterogeneity of not only data and technologies but also of organisations and intentions is a serious impediment to the development of holistic and ubiquitous smart cities. It is primordial to apply solution both technical and organisational to release its full potential. On that matter, great efforts for standardisation via open standards have been witnessed to achieve interoperability [31], [202], [319], [320], [635]. Those are essential to coordinate innovations via a common framework on top of which smart services can be designed. However, some argue that traditional standardisation may not be enough considering the fast growing and complex nature of the field [247]. The low penetration of the semantic web in the industry demonstrates a lack of consideration toward such technologies. This is regrettable as coupling such solution with open standards has demonstrated great performance in some research projects [27], [331], [367], [614], [622].

Additionally, early mentioned earlier, collaborations between organisations is essential. The domain would greatly benefit from the application of collaborative network theory. An important point to stress is the creation of public-private partnerships. Those cooperations must be constructed around new business models that are profitable for investors without being costly for the target users. Public institutions must provide good incentives to private organisations so that they do not develop dispatched and uncoordinated solutions. This task is all the more difficult that competition could play a restrictive role in the pursuit of such collaboration.

Another conflictual point comes with data ownership. It is important to stress that data belong to measurements objects. Therefore, citizens have sovereignty over the data they produce. This is not so evident in a context where IoTs' investments are shared with private organisations. Regulative frameworks are important to ensure data protection and that citizen's privacy is not violated. Such frameworks should comfort people to share their data safely and bring trust among all the parties involved. This includes securing, anonymizing the data and providing user-friendly open platforms to visualise and collect data. Beyond the regulative constraints, private and public organisations have the ethical responsibility to use citizen information with care and for genuine purposes.

On this last point, some experts warn on the risk of data and smart cities to support a neo-liberal economy driven by profit [392], [393]. They argue that recent solutions developed in the domain were mostly business-led, entrepreneurial or corporate urban development [392]. It is important that the IoT and smart cities serve people interests first, undistinguishably of their social status, rather than capitalising at their expense. The recent scandal over Facebook private data harvested by Cambridge Analytica to target political advertising is a perfect example of privacy violation for political and lucrative purposes [671]. Consistent regulations must be instore without exploitable loopholes that could allow dishonest purposes. Furthermore, the IoT provides the possibility for a ubiquitous control and surveillance which would violate fundamental human rights [18]. An example has been the mass surveillance of the phone data of millions of Americans by the U.S. intelligence agencies that was arguably morale [672]. There is a risk of political forces using such technologies to form undemocratic and authoritarian regimes and apply a never-seen-before pressure on their population. Consequently, if such approach presents incredible benefits, it can be harmful and safeguards must be applied to ensure such deviating purposes are unreachable. Finally, there are legitimate concerns on the creation of a technocratic system of governance of the city [18]. If smart solutions such as the one developed in this research have great potential to help decision making, they have no vocation to supplant human inputs. The opposite would be reductionist and give too much importance to data and their analysis. The threat is to create a data-dependent system with a too great influence of data experts.

7.2 ANSWERS TO THE RESEARCH QUESTIONS

RQ1. What are the issues that face the different stakeholders of an urban system toward sustainability assessment?

This question is mostly addressed with the systematic literature review of the existing USA schemes presented in Chapter 2. If the schemes are complete and well designed, covering many aspects that do sustainability, there are still certain gaps that can be filled. The review has identified a set of issues relating to their use (Section 2.1.3). Those frameworks are mostly used for design purposes with only 4 that considered existing urban systems. They do not go beyond a construction-oriented perspective while many stakeholders could benefit from an operation-oriented approach. Consequently, the design-oriented purposes of the schemes result in a biases consideration of environmental aspects and costs at the expense of socio-economical aspects. This witnesses a technocratic approach to tackle sustainability where technological fixes with economic outcomes are promoted to overcome challenges. Furthermore, in such context, there is no real need for dynamic systems to evaluate sustainability. However, shifting the vision toward urban areas in operation would require the inclusion of dynamism within the assessment. Indeed, neighbourhoods are complex, dynamic systems which requirements change with time. Therefore, stakeholders that desire to operate such assessment needs the support of frameworks that adapt themselves over time. The issues of adaptability can also be considered over location and culture. The literature review has demonstrated that requirements change with places and cultures and that the USA needs to adapt itself. They do so by stressing certain aspects more than others via their weighing system. However, in most of the cases, this adaptability is done in an ad-hoc manner with no to little room for automated mechanisms that could help in this task. If an ad-hoc approach is reliable and consistent, stakeholder could benefit from support systems that help in selecting the aspects to stress or the benchmark considered.

Finally, the last issue identified was the lack of consensus on the definition of sustainability. Indeed, the indicators' considerations and weights within the frameworks are disparate. Such observation witnesses a lack of common agreement on how to tackle sustainable development. If broad concepts such as the three dimension of sustainability and core themes reached an agreement among experts, opinions are divided into more detailed aspects.

RQ2. How an effective urban sustainability assessment can help the different parties of a city in their decision making?

Following the outcome of the literature review, developing a sustainability assessment framework for cities in operation could greatly help to support decisions of city operators and ultimately improve citizen quality of life through a neighbourhood entire lifecycle. This is especially important as researches have observed the positive influence of good maintenance and operation on facilities' performances and lifespan. Such approach confers the assessment with a real decision support function, going beyond a simple certification award as it has been observed. Decision supports systems have been observed and investigated in section 2.5 whether in the literature or through the projects involved in the action research approach. A common grounding in those projects is the necessity to present stakeholders with real or near real-time information so that they take the most efficient and relevant decisions, especially over quick changing aspects that require a fast demand-response. Moreover, they should be given analytical models that enable them to foresee possible changes and requirements at the urban level and prevent incoming issues. This will require the use of technologies able to capture information in real-time to calculate sustainability KPIs as well as more or less sophisticated forecasting models to predict future values (section 2.2, 2.5.2 & 4.3). Experts have been approached to participate in a DELPHI consultation that would set validity over the feasibility of a real-time urban sustainability assessment (section 4.2 & 4.4). The most important points stressed were the overall relevance, the ability to capture aspects in real-time, the ability to improve those aspects via operational and managerial efforts and recommendations on log frequency and forecast horizons. A selected few 17 indicators, mostly energy-related, were subject to the survey which validated them as relevant for the study, possibly improvable via managerial actions but that some may be costly to implement. In addition, those indicators were most likely considered as in real-time with hourly to daily measurements and forecasts up to a day/week were considered. Those questions remain open for the other indicators that compose the new scheme created, especially for those aspects quasi-static and/or with a great measurement complexity.

RQ3. How can sustainability assessment leverage on the smart city paradigm, specifically ICTs and the IoT?

Sustainability covers several domains which confer the field with an increased complexity in comparison to domain-specific decision support systems. This will require a large amount of data supported by a vast sensing network. The research has highlighted in section 2.2 the great potential that presents the smart city paradigm for such approach leveraging on key technologies that are the IoT, BIM, GIS, data mining and crowdsourcing. Evidences have been given in section 4.3 to support the measurement of every indicator addressed. Despite a relative implementation readiness (depending on the technology), there are strong groundings for further developments and deployments and great expectation over those. With the growing interest in smart solution for homes, wearable, health, surveillance, energy provision etc, the IoT presents a wide range of sensing devices able to capture multiple physical phenomena. The ability to measure energy, weight, volumes, noise, light, concentration, velocities etc in real-time is no longer questioned. The IoT goes along 4D BIM and GIS models with an increasing level of details and penetration within the industry where as-is urban space and facilities can be accurately described. Those technologies constitute a valuable source of knowledge for urban design and facilities management with the identification of objects, actors, spaces' functions, geometries etc. Their development is supported by remote sensing technologies such as LiDAR, high resolution satellite or aerial imagery that have shown great improvement in the last few years. Finally, the Internet has enabled the active and/or passive provision and collection of information by citizens and governmental institutions. The open access to such information resource is prised for the determination of socio-economic aspects. Indeed, crowdsourcing and data mining enable to capture extended qualitative information on the citizens' quality of life, concerns and satisfaction over their urban environment.

The implementation of the real-time USA framework with its web application platform has demonstrated the feasibility of using multiple sensors in the real-time determination of sustainability indicators over an entire neighbourhood. The system allowed its user to retrieve information about KPIs and provided 24h to 7 days forecasts.

RQ4. How can semantic web technologies unify heterogeneous data resources for holistic services and applications?

Despite being promising, the smart city paradigm and its technologies do not currently meet their full potential (section 2.2.3 & 2.3.3). Indeed, most (if not all) smart solutions projects studied were presented with heterogeneous information and systems. This diversity prevents the creation of holistic and ubiquitous systems. The integrity of the structure is ensured in an ad-hoc way where experts developed protocols, follows standard and coordinate efforts to homogenise the information. However, if such approach seems suitable for closed systems such as a single building or neighbourhood, scaling to a broader scale seems difficult. On that topic, some researchers have raised concerns about the "traditional" implementation of standards for information systems interoperability considering the plurality of solutions and the fast growing pace of the domain. In that context, the research has demonstrated the use of semantic web technologies as a valuable option to tackle information systems interoperability (section 2.3 & 5.2.4). The technology enhances information semantic, grounding for a common understanding of each piece of information not only by humans but also by machines. Therefore, it enables a seamless M2M communication above heterogeneous data and system. In this research, an Ontology-Based Data Access has been deployed based on the development of an USA ontology presented in chapter 5. The application was able to link data contained in BIM models with data produced by sensors, showcasing the interoperability of disparate sources of information. In such approach, the user is able to query the system using BIM constructs and still retrieve sensors information in reasonable computation times (chapter 6). Moreover, semantic web technologies allow, via the creation of rules and the use of reasoning engines, to infer extra pieces of information which would be difficulty discovered with other technologies. Consequently, beyond the acquired interoperability, such technologies brings intelligence via a greater understanding by machines of the data.

RQ5. How technological, human and financial assets relate to such service provision approach in the smart city paradigm?

If the readiness of real-time urban sustainability assessment framework supported by the IoT and semantic web technologies is promising, concerns are raised toward its reception and implementation on the larger scale by the industry. Observations from projects and the literature review in section 2.4 have led the author to advocate for the implementation of collaborative networks for the deployment of such type of solutions. Indeed, they will most likely involve various technological, human and financial assets and their success will then rely on the ability of the diverse entities to cooperate. In the current neo-liberal and competitive economy where information intelligence is a key enabler for performance and profit, such type of collaboration is not ensured. Consequently, the creation of public-private partnerships is favoured for the elaboration of such solutions with new business models bringing new incentives for private organisations than direct profit. Governmental institutions are central actors to coordinate such innovative developments and to ensure their success. They have the responsibility to provide regulation and create partnerships that enable it while protecting citizens' privacy and freedom. Finally, the last point to pressure is the unilateral sovereignty of citizens and/or organisations over the data they produce. In that regards, it is essential to include citizens within the loop and to increase their participation in policies making and services development via the collection of feedbacks and the deployment of open platforms. A participatory approach along with complete transparency on the data uses will produce a greater understanding of the field and bring trust of all the stakeholders.

7.3 LIMITATIONS AND FUTURE WORK

7.3.1 USA scheme validation

The pilot versions of the DELPHI consultation have witnessed in a low participation with a response rate below 15%. Such a low response rate implies substantial efforts to fulfil an objective of collecting between 30 and 50 responses from experts. The principal reason mentioned by the expert was the length of the survey with its 193 indicators to be reviewed. Efforts have been made to shorten the survey by grouping KPIs based on similar observations and by designing shortcuts to enable to jump certain sections. This has lowered the survey's objects from 193 to 71. Despite those efforts, only 6 responses out of 57 participants have been collected. The survey remained too long to be answered on a voluntary basis. A literature review on the topic of incentives for voluntary experts' surveys has demonstrated that monetary incentives would not work in increasing participation as experts are more driven by the subjects' interest than profit. Consequently, the study would have benefited from a well-defined research consortium where the experts would have more actively participated in the research outcome. The voluntary aspect of the survey resulted in that only its length and the efforts required would have an impact on the participation rate. The choice of addressing a selected few indicators have therefore been made on the basis of the future implementation plan. It is regrettable that the USA scheme has not been validated entirely; especially around singular indicators whose technological readiness, in real-time, is still questionable (e.g. remote sensing urban objects detection, social media data mining, crowdsource solutions etc).

Consequently, future efforts should be made for the validation of the entire USA scheme with greater incentives for the experts to participate. This includes the definition of a research project and consortium with different domains' experts and the release of domain-specific surveys, each containing a subset of the entire USA scheme. Once the scheme validated, additional efforts should be made to adapt the USA ontology to the scheme by simply removing the indicators rejected and their related axioms.

7.3.2 ONTOP maturity

Another limitation comes with the maturity of ONTOP for the OBDA development. Indeed, ONTOP rests on the OWL 2 QL entailment which limits the expressivity of the language in order to gain in performance. There is a set of restriction that stops reasoner from inferring information through transitive properties, existential and universal quantification, cardinality, disjunction etc. This forces the user to design more complex mappings to gain in expressivity. In addition, the SPARQL end-point integrated with ONTOP has restrictions as well which limits its ability to answer certain queries. As such, construct such as aggregates functions, cast functions, custom datatypes are not included. Same goes for geospatial functions that have been the subject of development of ONTOP-spatial, a branch of ONTOP that no longer seems to be fully active. Consequently, in term of maturity, ONTOP is still at a relatively early stage and further development should be done to overcome current limitations and deliver increasingly performant smart solutions.

The technology is promising and the ONTOP development greatly active. Its base community is increasing which prefigures further improvement in the near future. The industry and governmental institutions should support such initiatives by developing a clear agenda promoting semantic web technologies and OBDAs.

7.3.3 BIM and CityGML instantiation

One last major issue was the creation of RDF BIM and cityGML models. Indeed, despite the provision of an open source IFC-to-RDF converter, the conversion of complex BIM models into ifcOWL remains problematic. The output RDF files of the BIM model was too large for an in-memory use in Protégé and some inconsistency produced the impossibility to store it in an Apache Jena Triple Store. After contacting the author of the converter, it appeared that a program to produce simpler ifcOWL files was still a work in progress and its efficiency remained to be tested. Same goes for cityGML models where no conversion programs were found. Overall, the immaturity of the conversion tools has led to convert models from scratch in Protégé with a low level of details (e.g. definition of rooms, buildings and certain appliances and objects). Further development of the USA framework would benefit from the use of well detailed RDF BIM and cityGML models. Efforts on the creation of conversion tools should continue and BIM and GIS software should include the possibility to export model in RDF OWL format.

On a more general note, more should be done in various domains related to the city and its environment for the production and delivery of RDF data. Experts and institutions should stress the importance of including semantic web technologies within the ICT agenda for a greater penetration of the domain.

7.3.4 Decision support system

Decision support systems are the part of information systems that help and improve decision making [673]. The field has been strongly influenced by the work of Herbert Simon, Nobel Prize for his theory of decision making in 1978, and his model of decision-making process consisted of three phases: intelligence, design and choice [673]–[675]. The main features of DSS are: (a) an improved access to data; (b) facilitated analysis (through automated intelligence); and (c) a greater communication (meaningful and practical results presented in a user-friendly manner) [676]. Additionally, DSSs potentially involve improved strategic advantage; reduced time consumption; smarter response; enhanced consistency; worker empowerment; reduced costs; greater innovation; and higher retention [676], [677].

Moreover, DSSs are increasingly relevant within the smart city paradigm as they leverage IoT and Big Data. Indeed, datasets are significantly larger and with them the possibility of analysing the data stored [421]. It has the potential to enhance knowledge and to help to take informed decisions. In their survey toward Big Data and decision making for the business sector, the Economist Intelligence Unit shows that big data has already improved organisations’ performance by 25% to 30% and will continue over the next years[243].

In this perspective and as introduced in section 2.5, DSS theory should be explored thoroughly for the future development of the USA framework and tool. Efforts must stress core requirements from various stakeholders and go beyond forecasting by integrating additional features to improve the data analysis and the decision making (e.g. alerts, fault detection etc). This would allow the framework to go beyond simple scoring or assessing and to become an essential ally in the decision making at the urban level.

7.3.5 The USA Framework Validation

Chapter 6 presented an exhaustive list of checks and verifications, showcasing the feasibility of a real-time urban sustainability assessment application supported by the IoT and semantic web technologies. However, such application still requires some development tasks to ensure its full viability as a real world application. Future work should focus on the finalisation of the application and its deployment among partners and the public in order to proceed to a complete validation of the service. The application performance should be studied and feedbacks from users taken in consideration. Such iterative task will require a substantial amount of resources and time, which is why it exceeds the scope of this research.

7.4 CLOSING REMARK

This thesis has investigated the feasibility of a real-time urban sustainability assessment for neighbourhoods in operation. A new scheme, ontology and web application software have been designed constituting the new Urban Sustainability Assessment Framework. Its implementation demonstrated and discovered the value of ICTs and semantic web technologies for such type of smart solution, especially relevant within the smart city movement. Great effort remains to be done both technically and organisationally for the complete development of a holistic and ubiquitous information system to support sustainability assessment. However, this research presented a proof of concept that contributed to the body of knowledge of the domain by proposing valuable artefacts and leads to be considered.

References

[1] C. Martinez-Fernandez, I. Audirac, S. Fol, and E. Cunningham-Sabot, “Shrinking Cities: Urban Challenges of Globalization,” *Int. J. Urban Reg. Res.*, vol. 36, no. 2, pp. 213–225, 2012.

[2] R. Maunier, “The Definition of the City,” *Am. J. Sociol.*, vol. 15, no. 4, pp. 536–548, 1910.

[3] United Nations, “World Urbanization Prospects - Population Division - United Nations,” *DESA, Population Division*, 2018. [Online]. Available: https://population.un.org/wup/Maps/. [Accessed: 06-Feb-2019].

[4] European Environ Agency, *The impacts of urban sprawl*, no. 10. 2006.

[5] United Nations, “World Urbanization Prospects: The 2018 Revision,” 2018.

[6] Montgomery Mark R, “The Urban Transformation of the Developing World,” *Science*, 2008. [Online]. Available: http://science.sciencemag.org/. [Accessed: 05-Feb-2019].

[7] Y. Wei, C. Huang, P. T. I. Lam, and Z. Yuan, “Sustainable urban development: A review on urban carrying capacity assessment,” *Habitat Int.*, vol. 46, pp. 64–71, 2015.

[8] BRE, “BREEAM International New Construction 2016.” 2016.

[9] USGBC, “LEED v4 - guide.” p. 39, 2013.

[10] CASBEE, “CASBEE - Tools For Building Scale,” 2018. [Online]. Available: http://www.ibec.or.jp/CASBEE/english/toolsE_building.htm. [Accessed: 01-Aug-2018].

[11] HQE, “HQE international : certification HQE à l’international,” 2018. [Online]. Available: https://www.behqe.com/fr. [Accessed: 01-Aug-2018].

[12] DGNB, “DGNB System - Certificate for sustainable and green building,” 2017. [Online]. Available: https://www.dgnb-system.de/en/. [Accessed: 01-Aug-2018].

[13] J. Gil and J. P. Duarte, “Tools for evaluating the sustainability of urban design: a review,” *Urban Des. Plan.*, vol. 166, no. DP6, pp. 311–325, 2013.

[14] R. F. M. Ameen, M. Mourshed, and H. Li, “A critical review of environmental assessment tools for sustainable urban design,” *Environ. Impact Assess. Rev.*, vol. 55, pp. 110–125, 2015.

[15] A. Welch, K. Benfield, and M. Raimi, “A Citizen ’ s Guide to LEED for Neighborhood Development : How to Tell if Development is Smart and Green.” pp. 2–44, 2011.

[16] BRE Global Limited, “BREEAM Communities. Technical Manual,” 2012.

[17] Institute for Building Environment and Energy Conservation (IBEC), “CASBEE for Urban development. Comprehensive Assessment System for Built Environment

Efficiency. Technical Manual (2014 Edition),” p. 98, 2014.

[18] R. Kitchin, “The real-time city? Big data and smart urbanism,” *GeoJournal*, vol. 79, no. 1, pp. 1–14, 2014.

[19] A. Meijer, M. Pedro, and R. Bolívar, “Governing the smart city: a review of the literature on smart urban governance,” *Int. Rev. Adm. Sci.*, vol. 82, no. 2, pp. 392–408, 2016.

[20] Z. K. A. Mohamed and E. S. A. & Ahmed, “Internet of Things Applications , Challenges and Related Future Technologies,” *WSN World Sci. News*, vol. 62, no. 2, pp. 126–148, 2017.

[21] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, “Internet of Things (IoT): A vision, architectural elements, and future directions,” *Futur. Gener. Comput. Syst.*, vol. 29, no. 7, pp. 1645–1660, Sep. 2013.

[22] e-Madina, “Smart City Value Chain - White Paper,” 2016.

[23] D. Remund and D. D. Aikat, “Drowning in Data: A Review of Information Overload within Organizations and the Viability of Strategic Communication Principles,” in *Information Overload*, Hoboken, NJ, USA: John Wiley & Sons, Inc., 2012, pp. 231–250.

[24] A. Sharifi and A. Murayama, “A critical review of seven selected neighborhood sustainability assessment tools,” *Environ. Impact Assess. Rev.*, vol. 38, pp. 73–87, 2013.

[25] L. Sullivan, Y. Rydin, and C. Buchanan, “Neighbourhood Sustainability Frameworks-A Literature Review,” 2014.

[26] C. Kuster, Y. Rezgui, J.-L. Hippolyte, and M. Mourshed, “Total life cycle and near real time environmental assessment approach : an application to district and urban environment ?,” in *eWork and eBusiness in Architecture, Engineering and Construction: ECPPM 2016,* 2016.

[27] S. Howell, Y. Rezgui, and T. Beach, “Integrating building and urban semantics to empower smart water solutions,” *Autom. Constr.*, vol. 81, pp. 434–448, 2017.

[28] I. P. Žarko, K. Pripužic, and M. Serrano, *Interoperability and Open-Source Solutions for the Internet of Things*, vol. 10218. 2017.

[29] A. Kazmi, Z. Jan, A. Zappa, and M. Serrano, “Overcoming the Heterogeneity in the Internet of Things for Smart Cities,” vol. 10218, 2017, pp. 20–35.

[30] R. Minerva, A. Biru, and D. Rotondi, “Towards a definition of the Internet of Things.” Telecom Italia, 2015.

[31] European Commission, “Horizon 2020 -Work Programme 2018-2020 Information and Communication Technologies.” 2018.

[32] J. W. Creswell, “Philosophical Assumptions and Interpretive Frameworks,” in *Qualitative Inquiry & Research Design. Chosing Among Five Approaches*, 2013, pp. 273–301.

[33] J. Jacobs, *The Death and Life of Great American Cities*, vol. 71. 1993.

[34] BRE, “In-Use – BREEAM.” [Online]. Available: https://www.breeam.com/discover/technical-standards/breeam-in-use/. [Accessed: 01-Aug-2018].

[35] BRE Global Limited, *BREEAM UK Refurbishment and Fit-out*. 2016.

[36] A. Brits, M. Burke, and T. Li, “Improved modelling for urban sustainability assessment and strategic planning: local government planner and modeller perspectives on the key challenges,” *Aust. Plan.*, vol. 51, no. 1, pp. 76–86, Jan. 2014.

[37] U. Berardi, “Beyond Sustainability Assessment Systems: Upgrading Topics by Enlarging The Scale of Assessment,” *Int. J. Sustain. Build. Technol. Urban Dev.*, vol. 2, no. 4, pp. 276–282, 2011.

[38] R. A. Bohne, O. J. Klakegg, and O. Lædre, “Evaluating Sustainability of Building Projects in Urban Planning,” *Procedia Econ. Financ.*, vol. 21, pp. 306–312, 2015.

[39] S. Curwell and I. Cooper, “The implications of urban sustainability,” *Build. Res. Inf.*, vol. 26, no. 1, pp. 17–28, 1998.

[40] Criterion Planners, “A Global Survey of Urban Sustainability Rating Tools,” 2014.

[41] BREEAM, “BREEAM Projects,” 2018. [Online]. Available: https://tools.breeam.com/projects/explore/buildings.jsp?sectionid=10154&projectType=&rating=&certNo=&buildingName=&client=&developer=&certBody=&assessor=&location=&countryID=0&partid=10023&Submit=Search. [Accessed: 02-Aug-2018].

[42] USGBC, “Projects | U.S. Green Building Council,” *USGBC*, 2019. [Online]. Available: https://www.usgbc.org/projects. [Accessed: 02-Aug-2018].

[43] CASBEE, “CASBEE Certification System Projects,” 2018. [Online]. Available: http://www.ibec.or.jp/CASBEE/english/certificationE.htm. [Accessed: 02-Aug-2018].

[44] CRETIVEA, “Plateforme CERTIVEA for HQE Projects,” 2018. [Online]. Available: http://certibox.certivea.fr/. [Accessed: 02-Aug-2018].

[45] DGNB, “DGNB registered and certified projects,” *DGNB-System*, 2018. [Online]. Available: https://www.dgnb-system.de/en/projects/index.php?filter_Freitextsuche=&filter_Land=&filter_Bundesland=&filter_Standort=&filter_Jahr=&filter_Zertifizierungsart=&filter_Nutzungsprofil=&filter_Zertifiziert_von_1=DGNB+%28International+Version%29&filter_Verlie. [Accessed: 02-Aug-2018].

[46] A. Haapio, “Towards sustainable urban communities,” *Environ. Impact Assess. Rev.*, vol. 32, no. 1, pp. 165–169, 2012.

[47] M. Orova and A. Reith, “Comparison and evaluation of neighbourhood sustainability assessment systems,” *PLEA 2013 Sustain. Archit. a Renew. Futur.*, no. September, 2013.

[48] A. Sharifi and A. Murayama, “Neighborhood sustainability assessment in action: Cross-evaluation of three assessment systems and their cases from the US, the UK, and Japan,” *Build. Environ.*, vol. 72, pp. 243–258, 2014.

[49] Criterion Planner, “Transformative Tools,” 2018. [Online]. Available:

http://www.transformativetools.org/. [Accessed: 05-Aug-2018].

[50] Brede, "Berde for New construction: Clustered Residential Development," 2013.

[51] CASBEE, "CASBEE for Cities. Pilot version for worldwise use," 2015.

[52] CEEQUAL Ltd., "Assessment Manual for UK & Ireland Projects Version 5," no. August, 2012.

[53] American Planning Association, "Sustaining Places Practices for Comprehensive Planning," 2015.

[54] Green Building Council Denmark, "An introduction to DGNB - Ensure the quality of your sustainable buildings in planning, construction, and operation.," 2012.

[55] Portland Sustainability Institute, "The EcoDistricts Toolkit : Action Guide," no. April, pp. 1–8, 2012.

[56] Portland Sustainability Institute, "The EcoDistricts Toolkit : Assessment. Prioritizing Projects in an EcoDistrict," 2011.

[57] Ministère de l'égalité des territoires et du logement, "Dossier de labellisation-EcoQuartier," pp. 1–8, 2011.

[58] N. Zhou, G. He, C. Williams, and D. Fridley, "ELITE cities: A low-carbon eco-city evaluation tool for China," *Ecol. Indic.*, vol. 48, pp. 448–456, 2015.

[59] D. Belzer *et al.*, "2015 Enterprise Green Communities Criteria," 2015.

[60] EnviroDevelopment, "EnviroDevelopment National Technical Standards Version 2." 2013.

[61] Institute for Sustainable Infrastructure and Zofnass Program for Sustainable Infrastructure, "Envision Rating System for Sustainable Infrastructure," p. 168, 2015.

[62] Green building index, "GBI Assessment Criteria for Township (Version 1.01)," no. September, pp. 0–57, 2011.

[63] Golf Organisation for Research & Development, "Global Sustainability Assessment System (GSAS) Sustainability Promises," 2014.

[64] Gold Standard, "Gold Standard Cities." [Online]. Available: https://www.goldstandard.org/. [Accessed: 22-Feb-2018].

[65] BCA, "BCA Green Mark for Districts," pp. 1–24, 2013.

[66] BCA, "BCA Green Mark for Infrastructure," pp. 1–24, 2013.

[67] Green Building Council of Australia, "Green Star - Communities," 2012.

[68] BeHQE, "HQE international : certification HQE à l'international." [Online]. Available: http://www.behqe.com/fr. [Accessed: 22-Feb-2018].

[69] IGBC, "IGBC Green Townships (For Townships and Large Developments)," no. November, 2010.

[70] USGBC, "LEED v 4 for NEIGHBORHOOD DEVELOPMENT," 2014.

[71] The International Living Future Institute, "Living Community Challenge 1.0," pp. 1–62, 2014.

[72] Home Innovation Research Labs, "NGBS Green Certification for Land Development." [Online]. Available: http://www.homeinnovation.com/services/certification/green_homes/land_development_certification. [Accessed: 22-Feb-2018].

[73] Abu Dhabi Urban Planning Council, "The Pearl Rating System for Estidama Community Rating System." 2010.

[74] Green Business Certification Inc., "Peer - Rating System Version 1.1." 2016.

[75] STAR Communities, "Technical Guide to the STAR Community Rating System," p. 133, 2015.

[76] U. Ranhagen and K. Groth, *The SymbioCity Approach*. 2012.

[77] Realdania By, "Tool for Sustainable Urban Development." 2013.

[78] N. Munier, *Multicriteria Environmental Assessment A Practical Guide*. Dordrecht: Kluwer Academic Publishers, 2004.

[79] A. Zeinal Hamedani and F. Huber, "A comparative study of DGNB, LEED and BREEAM certificate systems in urban sustainability," *WIT Trans. Ecol. Environ.*, vol. 155, no. August, pp. 121–132, 2011.

[80] B. R. Keeble, "The Brundtland Report: 'Our Common Future,'" *Medicine and War*, vol. 4, no. 1. pp. 17–25, 1988.

[81] G. A. Tanguay, J. Rajaonson, J. F. Lefebvre, and P. Lanoie, "Measuring the sustainability of cities: An analysis of the use of local indicators," *Ecol. Indic.*, vol. 10, no. 2, pp. 407–418, 2010.

[82] G. T. Doran, "There's a S.M.A.R.T. way to write managements's goals and objectives.," *Manage. Rev.*, vol. 70, no. 11, pp. 35–36, 1981.

[83] L. Huang, J. Wu, and L. Yan, "Defining and measuring urban sustainability: a review of indicators," *Landscape Ecology*, vol. 30, no. 7. pp. 1175–1193, 2015.

[84] R. C. Retzlaff, "The use of Leed in Planning and Development Regulation: An Exploratory Analysis," *J. Plan. Educ. Res.*, pp. 1–11, 2009.

[85] K. M. Davidson, J. Kellett, L. Wilson, and S. Pullen, "Assessing urban sustainability from a social democratic perspective: a thematic approach," *Local Environ.*, vol. 17, no. 1, pp. 57–73, 2012.

[86] R. K. Singh, H. R. Murty, S. K. Gupta, and A. K. Dikshit, "An overview of sustainability assessment methodologies," *Ecol. Indic.*, vol. 15, no. 1, pp. 281–299, 2012.

[87] J. Rotmans, M. Van Asselt, and P. Vellinga, "An integrated planning tool for sustainable cities," *Environ. Impact Assess. Rev.*, vol. 20, no. 3, pp. 265–276, 2000.

[88] European Committee for Standardization, "CEN/TC 350 Business Plan," 2013.

[89] J. Joshi, "Sustainability of Heritage Neighborhoods in Kathmandu," *Int. J. Soc. Sci.*

Humanit., vol. 6, no. 3, pp. 221–225, Mar. 2016.

[90] J. Keddie, “Negotiating urban change in gentrifying London : Experiences of long-term residents and early gentrifiers in Bermondsey,” no. April, 2014.

[91] National Research Council, *Fourth Dimension in Building: Strategies for Avoiding Obsolescence*. Washington, D.C.: National Academies Press, 1993.

[92] Pacific Northwest national Laboratory and U.S department of Energy, “PNNL Building Systems - Fault Detection & Diagnostics,” 2013. [Online]. Available: https://buildingsystems.pnnl.gov/fdd/fdd.stm. [Accessed: 02-Oct-2018].

[93] S. R. Chandratilake and W. P. S. Dias, “Sustainability rating systems for buildings: Comparisons and correlations,” *Energy*, vol. 59, pp. 22–28, 2013.

[94] D. Childers, M. Cadenasso, J. Grove, V. Marshall, B. McGrath, and S. Pickett, “An Ecology for Cities: A Transformational Nexus of Design and Ecology to Advance Climate Change Resilience and Urban Sustainability,” *Sustainability*, vol. 7, no. 4, pp. 3774–3791, Mar. 2015.

[95] J. Ahern, “Novel Urban Ecosystems: Concepts, Definitions and a Strategy To Support Urban Sustainability and Resilience,” *Landsc. Archit. Reg. Plan. Ser. 66*, pp. 10–21, 2016.

[96] J. Wu, “Making the case for landscape ecology: an effective approach to urban sustainability,” *Landsc. J.*, vol. 27, pp. 41–50, 2008.

[97] C. J. Tanner *et al.*, “Urban ecology: Advancing science and society,” *Front. Ecol. Environ.*, vol. 12, no. 10, pp. 574–581, 2014.

[98] W. H. Dutton, J. G. Blumler, and K. L. Kraemer, *Wired cities: Shaping the future of communications*. 1987.

[99] S. Graham and S. Marvin, “Planning cyber-cities? Integrating telecommunications into urban planning,” *Town Plan. Rev.*, vol. 70, no. 1, pp. 89–114, 1999.

[100] T. Ishida and K. Isbister, *Digital cities: Technologies, experiences, and future perspectives*. Springer, 2000.

[101] N. Komninos, “Intelligent Cities: Innovation, Knowledge Systems and Digital Spaces,” in *Routledge: Knowledge Systems and Digital Spaces*, 2002.

[102] F. J. Carrillo, *Knowledge cities: Approaches, experiences and perspectives*. Routledge, 2005.

[103] Leslie Brokaw, “Six Lessons From Amsterdam’s Smart City Initiative,” 2016. [Online]. Available: https://sloanreview.mit.edu/article/six-lessons-from-amsterdams-smart-city-initiative/. [Accessed: 06-Jun-2018].

[104] M. Fitzgerald, “Data-Driven City Management: A Close Look at Amsterdam’s Smart City Initiative,” *MIT Sloan Manag. Rev.*, vol. 57, no. 4, p. n/a, 2016.

[105] Forbes, “Top Five Smart Cities,” 2015. [Online]. Available: https://www.forbes.com/pictures/54fe2629fcd7da7ddff45018/top-five-smart-cities/#21e59806b431. [Accessed: 07-Jun-2018].

[106] B. Buntz, “The Top 5 Smart Cities,” *May 18*, 2016. [Online]. Available: http://www.ioti.com/smart-cities/world-s-5-smartest-cities. [Accessed: 07-Jun-2018].

[107] Juniper Research, “Smart Cities — What’s in it for Citizens?,” *White Pap.*, 2018.

[108] J. McGrath, “What Makes Barcelona One of the Smartest Cities in the World? | Digital Trends,” *Digital Trends*, 2017. [Online]. Available: https://www.digitaltrends.com/home/barcelona-smart-city-technology/. [Accessed: 07-Jun-2018].

[109] T. Bakıcı, E. Almirall, and J. Wareham, “A Smart City Initiative: the Case of Barcelona,” *J. Knowl. Econ.*, vol. 4, no. 2, pp. 135–148, Jun. 2013.

[110] T. Sheu, “Smart City Barcelona: One City’s Quest For Innovation,” 2015. [Online]. Available: http://www.barcinno.com/smart-city-barcelona/. [Accessed: 07-Jun-2018].

[111] M. Gascó-Hernandez, “Building a Smart City: Lesson from Barcelona,” *Commun. ACM*, vol. 61, no. 4, pp. 50–57, Mar. 2018.

[112] J. Ancheta, “Ten Reasons Why Barcelona is a Smart City - VilaWeb,” *VilaWeb*, 2014. [Online]. Available: https://www.vilaweb.cat/noticia/4175829/20140226/ten-reasons-why-barcelona-is-smart-city.html. [Accessed: 07-Jun-2018].

[113] Francesca Bria, “Barcelona’s Strategy for Digital Sovereignty,” 2017. [Online]. Available: https://www.youtube.com/watch?v=YVvwsfpU51k. [Accessed: 07-Jun-2018].

[114] Smart Nation, “Strategic Projects and Enablers,” *Ministry of Communication and Information*, 2018. [Online]. Available: https://www.smartnation.sg/about/Smart-Nation. [Accessed: 07-Jun-2018].

[115] Singapore Ministry of Communication and Information, “Digital Readiness Blueprint.” 2017.

[116] Infocomm Media Development Authority, “Digital Economy Framework for Action,” *SG Digit.*, p. 53, 2018.

[117] GovTech Singapore, “Digital Government Blueprint - A Singapore Government that is Digital to the Core, and Serves with Heart,” 2018.

[118] J. Yeo, “Singapore and Smart Nation: Less privacy, greater sustainability?,” *ASEAN Today*, 2018. [Online]. Available: https://www.aseantoday.com/2016/05/singapore-and-smart-nation/. [Accessed: 07-Jun-2018].

[119] A. Aravindan and J. Geddie, “Singapore to test facial recognition on lampposts, stoking privacy fears,” *Reuters*, 2018. [Online]. Available: https://www.reuters.com/article/us-singapore-surveillance/singapore-to-test-facial-recognition-on-lampposts-stoking-privacy-fears-idUSKBN1HK0RV. [Accessed: 07-Jun-2018].

[120] L. Poon, “Singapore, City of Sensors,” *Citylab*, 2017. [Online]. Available: https://www.citylab.com/life/2017/04/singapore-city-of-sensors/523392/. [Accessed: 07-Jun-2018].

[121] Markets and Markets, “Smart Cities Market by Smart Transportation (Type, Solutions and Services), Smart Buildings (Type, Solutions and Services), Smart Utilities (Type, Solutions and Services), Smart Citizen Services, and Region - Global Forecast to 2023,” 2017. [Online]. Available: https://www.marketsandmarkets.com/Market-Reports/smart-cities-market-542.html. [Accessed: 07-Jun-2018].

[122] Grand View Research, “Smart Cities Market Size Worth $2.57 Trillion By 2025 | CAGR: 18.4%,” 2018. [Online]. Available: https://www.grandviewresearch.com/press-release/global-smart-cities-market. [Accessed: 07-Jun-2018].

[123] Ericsson, “Internet of Things forecast,” *Internet of Things forecast*, 2016. [Online]. Available: https://www.ericsson.com/en/mobility-report/internet-of-things-forecast. [Accessed: 07-Jun-2018].

[124] A. Danigelis, “Global Smart Cities Market To Top $2 Trillion by 2025, Analysts Say,” *Energy Manager Today*, 2018. [Online]. Available: https://www.energymanagertoday.com/global-smart-cities-market-frost-sullivan-0176058/. [Accessed: 07-Jun-2018].

[125] Z. Khan, A. Anjum, and S. L. Kiani, “Cloud based big data analytics for smart future cities,” *Proc. - 2013 IEEE/ACM 6th Int. Conf. Util. Cloud Comput. UCC 2013*, pp. 381–386, 2013.

[126] A. Caragliu, C. del Bo, and P. Nijkamp, “Smart cities in Europe,” *J. Urban Technol.*, vol. 18, no. 2, pp. 65–82, 2011.

[127] M. Hashmi, S. Hänninen, and K. Mäki, “Survey of smart grid concepts, architectures, and technological demonstrations worldwide,” *2011 IEEE PES Conf. Innov. Smart Grid Technol. Lat. Am. SGT LA 2011 - Conf. Proc.*, pp. 1–7, 2011.

[128] H. Adeli and X. Jiang, *Intelligent Infrastructure: Neural Networks, Wavelets, and Chaos Theory for Intelligent Transportation Systems and Smart Structures*. 2008.

[129] H. Demirkan, “A smart healthcare systems framework,” *IT Prof.*, vol. 15, no. 5, pp. 38–45, 2013.

[130] H. Chourabi *et al.*, “Understanding smart cities: An integrative framework,” in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2012, pp. 2289–2297.

[131] I. A. T. Hashem *et al.*, “The role of big data in smart city,” *Int. J. Inf. Manage.*, vol. 36, no. 5, pp. 748–758, Oct. 2016.

[132] E. Borgia, “The internet of things vision: Key features, applications and open issues,” *Comput. Commun.*, vol. 54, pp. 1–31, 2014.

[133] G. Hancke, B. Silva, and G. Hancke, Jr., “The Role of Advanced Sensing in Smart Cities,” *Sensors*, vol. 13, no. 1, pp. 393–425, Dec. 2012.

[134] R. Kadam, P. Mahamuni, and Y. Parikh, “Smart Home System,” *Int. J. Innov. Res. Adv. Eng.*, vol. 2, no. 1, pp. 2349–2163, 2015.

[135] T. Tahmid and E. Hossain, “Density based smart traffic control system using canny

edge detection algorithm for congregating traffic information," *3rd Int. Conf. Electr. Inf. Commun. Technol. EICT 2017*, vol. 2018–Janua, no. December, pp. 1–5, 2018.

[136] A. Khanna and R. Anand, "IoT based smart parking system," pp. 266–270, 2016.

[137] E. Aubry, T. Silverston, A. Lahmadi, and O. Festor, "CrowdOut: A mobile crowdsourcing service for road safety in digital cities," *2014 IEEE Int. Conf. Pervasive Comput. Commun. Work. PERCOM Work. 2014*, pp. 86–91, 2014.

[138] A. Papa, M. Mital, P. Pisano, and M. Del Giudice, "E-health and wellbeing monitoring using smart healthcare devices: An empirical investigation," *Technol. Forecast. Soc. Change*, Mar. 2018.

[139] M. Chen, Y. Ma, J. Song, C. F. Lai, and B. Hu, "Smart Clothing: Connecting Human with Clouds and Big Data for Sustainable Health Monitoring," *Mob. Networks Appl.*, vol. 21, no. 5, pp. 825–845, Oct. 2016.

[140] T. Lemlouma, S. Laborie, and P. Roose, "Toward a context-aware and automatic evaluation of elderly dependency in smart homes and cities," in *2013 IEEE 14th International Symposium on a World of Wireless, Mobile and Multimedia Networks, WoWMoM 2013*, 2013, no. September 2014, pp. 1–6.

[141] M. Sorek-Hamer, A. C. Just, and I. Kloog, "The Use of Satellite Remote Sensing in Epidemiological Studies," *Curr. Opin. Pediatr.*, vol. 28, no. 2, pp. 228–234, Apr. 2016.

[142] H. Lund *et al.*, "4th Generation District Heating (4GDH)," *Energy*, vol. 68, pp. 1–11, Apr. 2014.

[143] S. Howell, Y. Rezgui, J. L. Hippolyte, B. Jayan, and H. Li, "Towards the next generation of smart grids: Semantic and holonic multi-agent management of distributed energy resources," *Renew. Sustain. Energy Rev.*, vol. 77, no. November 2015, pp. 193–214, 2017.

[144] M. O. De Mues, A. Alvarez, A. Espinoza, and J. Garbajosa, "Towards a distributed intelligent ICT architecture for the smart grid," *IEEE Int. Conf. Ind. Informatics*, no. 1, pp. 745–749, 2011.

[145] L. Berrío and C. Zuluaga, "Concepts, standards and communication technologies in smart grid," *2012 IEEE 4th Colomb. Work. Circuits Syst. CWCAS 2012 - Conf. Proc.*, 2012.

[146] L. M. Camarinha-Matos, "Collaborative smart grids – A survey on trends," *Renew. Sustain. Energy Rev.*, vol. 65, pp. 283–294, 2016.

[147] and J. S. S. Chen, S. Song, L. Li, "Survey on smart grid technology," *(in Chinese). Power Syst. Technol.*, vol. 33, no. 8, pp. 1–7, 2009.

[148] A. Sadeghioon, N. Metje, D. Chapman, and C. Anthony, "SmartPipes: Smart Wireless Sensor Networks for Leak Detection in Water Pipelines," *J. Sens. Actuator Networks*, vol. 3, no. 1, pp. 64–78, 2014.

[149] Y. Zhang, S. M. O'Connor, G. W. van der Linden, A. Prakash, and J. P. Lynch, "SenStore: A Scalable Cyberinfrastructure Platform for Implementation of Data-to-Decision Frameworks for Infrastructure Health Management," *J. Comput. Civ. Eng.*, vol. 30, no. 5, p. 04016012, Sep. 2016.

[150] F. Bibi, C. Guillaume, N. Gontard, and B. Sorli, “A review: RFID technology having sensing aptitudes for food industry and their contribution to tracking and monitoring of food products,” *Trends in Food Science and Technology*, vol. 62. pp. 91–103, 2017.

[151] T. Reisbacka, H. Hamalainen, and J. Ikonen, “Automating construction project quality assurance with RFID- And mobile technologies,” *SoftCom 2008 16th Int. Conf. Software, Telecommuncations Comput. Networks*, pp. 223–227, 2008.

[152] B. Chowdhury and M. U. Chowdhury, “RFID-based real-time smart waste management system,” *2007 Australas. Telecommun. Networks Appl. Conf. ATNAC 2007*, pp. 175–180, 2008.

[153] S. Kumar and S. K. Singh, “Monitoring of pet animal in smart cities using animal biometrics,” *Futur. Gener. Comput. Syst.*, vol. 83, pp. 553–563, Jun. 2018.

[154] M. Ahmad, “4G and 5G wireless: How they are alike and how they differ,” 2015. [Online]. Available: http://www.androidauthority.com/4g-and-5g-wireless-how-they-are-alike-and-how-they-differ-615709/. [Accessed: 05-Jun-2018].

[155] L. Vinet and A. Zhedanov, *A “missing” family of classical orthogonal polynomials*, vol. 44, no. 8. 2011.

[156] E. Turban, *Decision support and expert systems: management support systems*. 1989.

[157] S. Yamamoto, S. Matsumoto, and M. Nakamura, “Using cloud technologies for large-scale house data in smart city,” *CloudCom 2012 - Proc. 2012 4th IEEE Int. Conf. Cloud Comput. Technol. Sci.*, pp. 141–148, 2012.

[158] B. Snaith, M. Hardy, and A. Walker, “Emergency ultrasound in the prehospital setting: The impact of environment on examination outcomes,” *Emerg. Med. J.*, vol. 28, no. 12, pp. 1063–1065, 2011.

[159] S. Arthur, H. Li, and R. Lark, “A collaborative unified computing platform for building information modelling (BIM),” in *IFIP Advances in Information and Communication Technology*, vol. 506, 2017, pp. 63–73.

[160] T. Beach, O. Rana, Y. Rezgui, and M. Parashar, “Governance Model for Cloud Computing in Building Information Management,” *IEEE Trans. Serv. Comput.*, vol. 8, no. 2, pp. 314–327, 2015.

[161] M. Mehdi, E. Curry, R. Sahay, W. Derguech, W. Qu, and S. Debloch, “Real-time Generation of Linked Sensor Data and Multidimensional Data Cubes for Smart Environments,” vol. 1, pp. 1–5, 2009.

[162] A. Vahid Dastjerdi, H. Gupta, R. N. Calheiros, S. K. Ghosh, and R. Buyya, “Fog Computing: Principles, Architectures, and Applications.”

[163] S. Hwang and L. Y. Liu, “BIM for Integration of Automated Real-Time Project Control Systems,” in *Construction Research Congress 2010*, 2010, pp. 509–517.

[164] H. Hajian and B. Becerik-Gerber, “A Research Outlook for Real-Time Project Information Management by Integrating Advanced Field Data Acquisition Systems and Building Information Modeling,” in *Computing in Civil Engineering (2009)*, 2009,

pp. 83–94.

[165] R. Volk, J. Stengel, and F. Schultmann, “Building Information Modeling (BIM) for existing buildings - Literature review and future needs,” *Autom. Constr.*, vol. 38, pp. 109–127, Mar. 2014.

[166] Q. Lu, J. Won, and J. C. P. Cheng, “A financial decision making framework for construction projects based on 5D Building Information Modeling (BIM),” *Int. J. Proj. Manag.*, vol. 34, no. 1, pp. 3–21, Jan. 2016.

[167] F. K. T. Cheung, J. Rihan, J. Tah, D. Duce, and E. Kurul, “Early stage multi-level cost estimation for schematic BIM models,” *Autom. Constr.*, vol. 27, pp. 67–77, Nov. 2012.

[168] P. Smith, “Project Cost Management with 5D BIM,” *Procedia - Soc. Behav. Sci.*, vol. 226, pp. 193–200, Jul. 2016.

[169] P. Yung and X. Wang, “A 6D CAD model for the automatic assessment of building sustainability,” *Int. J. Adv. Robot. Syst.*, vol. 11, no. 1, p. 131, Aug. 2014.

[170] F. Jalaei and A. Jrade, “Integrating building information modeling (BIM) and LEED system at the conceptual design stage of sustainable buildings,” *Sustain. Cities Soc.*, vol. 18, pp. 95–107, Nov. 2015.

[171] M. Bew and M. Richard, “Bew-Richards BIM maturity model,” *BuildingSMART Construct IT Autumn Members Meeting, Brighton*. BuildingSMART Construct IT Autumn Members Meeting, Brighton, 2008.

[172] HM Treasury and Cabinet Office, “Government Construction Strategy 2016-20,” 2016.

[173] A. Grilo and R. Jardim-Goncalves, “Value proposition on interoperability of BIM and collaborative working environments,” *Autom. Constr.*, vol. 19, no. 5, pp. 522–530, Aug. 2010.

[174] M. F. Muller, E. R. Loures, and O. C. Junior, “Interoperability Assessment for Building Information Modelling,” in *3rd International Conference on Mechatronics, Robotics and Automation (ICMRA 2015)* , 2015.

[175] C. Z. Li *et al.*, “Integrating RFID and BIM technologies for mitigating risks and improving schedule performance of prefabricated house construction,” *J. Clean. Prod.*, vol. 165, pp. 1048–1062, Nov. 2017.

[176] A. V. Barenji and M. Hashemipoor, “Real-Time Building Information Modeling (BIM) Synchronization Using Radio Frequency Identification Technology and Cloud Computing System,” *BIM-server*, vol. 10, no. 1, pp. 61–68, 2017.

[177] C. Wang and Y. Cho, “Automated As-is BIM Extraction for Sustainable Simulation of Built Environments.,” *Fifth Int. Conf. Constr. Eng. Proj. Manag.*, no. December 2017, 2013.

[178] J. Wang *et al.*, “Integrating BIM and LiDAR for Real-Time Construction Quality Control,” *J. Intell. Robot. Syst. Theory Appl.*, vol. 79, no. 3–4, pp. 417–432, 2015.

[179] C. Wang, Y. K. Cho, and C. Kim, “Automatic BIM component extraction from point

clouds of existing buildings for sustainability applications," *Autom. Constr.*, vol. 56, pp. 1–13, Aug. 2015.

[180] S. Murali, P. Speciale, M. R. Oswald, and M. Pollefeys, "Indoor Scan2BIM: Building information models of house interiors," in *IEEE International Conference on Intelligent Robots and Systems*, 2017, vol. 2017–Septe, pp. 6126–6133.

[181] Z. Shang and Z. Shen, "Real-time 3D Reconstruction on Construction Site using Visual SLAM and UAV," *2018 Constr. Res. Congr. Porc.*, 2017.

[182] Q. Lu, S. Lee, and L. Chen, "Image-driven fuzzy-based system to construct as-is IFC BIM objects," *Autom. Constr.*, vol. 92, pp. 68–87, Aug. 2018.

[183] C. Kropp, C. Koch, and M. König, "Interior construction state recognition with 4D BIM registered image sequences," *Autom. Constr.*, vol. 86, pp. 11–32, Feb. 2018.

[184] S. El-Omari and O. Moselhi, "Hybrid Methodology for Automated Collection of Data from Construction Site," in *Automation and Robotics in Construction ― Proceedings of the 24th International Symposium on Automation and Robotics in Construction*, 2017.

[185] S. El-Omari and O. Moselhi, "Integrating 3D laser scanning and photogrammetry for progress measurement of construction work," *Autom. Constr.*, vol. 18, no. 1, pp. 1–9, Dec. 2008.

[186] Caitlin Dempsey, "History of GIS," *GIS Lounge*, 2012. [Online]. Available: https://www.gislounge.com/history-of-gis/. [Accessed: 22-May-2018].

[187] J. Wright, T. M. Lillesand, and R. W. Kiefer, "Remote Sensing and Image Interpretation," *Geogr. J.*, vol. 146, no. 3, p. 448, 2006.

[188] I. R. Hegazy and M. R. Kaloop, "Monitoring urban growth and land use change detection with GIS and remote sensing techniques in Daqahlia governorate Egypt," *Int. J. Sustain. Built Environ.*, vol. 4, no. 1, pp. 117–124, Jun. 2015.

[189] D. Contreras, T. Blaschke, D. Tiede, and M. Jilge, "Monitoring recovery after earthquakes through the integration of remote sensing, GIS, and ground observations: The case of L'Aquila (Italy)," *Cartogr. Geogr. Inf. Sci.*, vol. 43, no. 2, pp. 115–133, Mar. 2016.

[190] G. M. Gandhi, S. Parthiban, N. Thummalu, and A. Christy, "Ndvi: Vegetation Change Detection Using Remote Sensing and Gis - A Case Study of Vellore District," in *Procedia Computer Science*, 2015, vol. 57, pp. 1199–1210.

[191] Michael Clancy, "High Resolution Aerial Imagery," *elecdata blog*, 2014. [Online]. Available: http://www.elecdata.com/blog/high-resolution-aerial-imagery/. [Accessed: 23-May-2018].

[192] G. Cheng and J. Han, "A survey on object detection in optical remote sensing images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 117. Elsevier, pp. 11–28, 01-Jul-2016.

[193] X. Bai, H. Zhang, and J. Zhou, "VHR object detection based on structural feature extraction and query expansion," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6508–6520, 2014.

[194] E. Li, J. Femiani, S. Xu, X. Zhang, and P. Wonka, "Robust rooftop extraction from visible band images using higher order CRF," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 8, pp. 4483–4495, 2015.

[195] J. Wang, J. Song, M. Chen, and Z. Yang, "Road network extraction: a neural-dynamic framework based on deep learning and a finite state machine," *Int. J. Remote Sens.*, vol. 36, no. 12, pp. 3144–3169, Jun. 2015.

[196] R. Margolis *et al.*, "Environmental Research Letters Using GIS-based methods and lidar data to estimate rooftop solar technical potential in US cities Rooftop solar photovoltaic potential in cities: how scalable are assessment approaches? Climate, air quality and human health ," *Environ. Res. Lett*, vol. 13, 2018.

[197] H. Mitasova, L. Mitas, and R. S. Harmon, "Simultaneous Spline Interpolation and Topographic Analysis for LIDAR Elevation Data: Methods for Open Source GIS," *IEEE Geosci. Remote Sens. Lett.*, vol. 2, no. 4, pp. 375–379, 2005.

[198] S. Shirowzhan and J. Trinder, "Building Classification from Lidar Data for Spatio-temporal Assessment of 3D Urban Developments," in *Procedia Engineering*, 2017, vol. 180, pp. 1453–1461.

[199] D. M. Campanaro, G. Landeschi, N. Dell'Unto, and A. M. Leander Touati, "3D GIS for cultural heritage restoration: A 'white box' workflow," *J. Cult. Herit.*, vol. 18, pp. 321–332, Mar. 2016.

[200] ISO, "ISO/PAS 17506:2012 Industrial Automation Systems and Integration—COLLADA Digital Asset Schema Specification for 3D Visualization of Industrial Data for 3D Visualization of Industrial Data," *International Organization for Standardization*, vol. 2012. 2012.

[201] Open Geospatial Consortium, "OGC Keyhole Markup Language 2.3." p. 266, 2015.

[202] OGC, "OGC® Geography Markup Language (GML) — Extended schemas and encoding rules," *OpenGIS Recomm. Pap.*, p. 595, 2010.

[203] P. Tymkow, M. Karpina, and A. Borkowski, "3D Gis for Flood Modelling in River Valleys," *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. XLI-B8, pp. 175–178, Jun. 2016.

[204] F. Biljecki, H. Ledoux, and J. Stoter, "An improved LOD specification for 3D building models," *Comput. Environ. Urban Syst.*, vol. 59, pp. 25–37, Sep. 2016.

[205] Y. Deng, J. C. P. Cheng, and C. Anumba, "Mapping between BIM and 3D GIS in different levels of detail using schema mediation and instance comparison," *Autom. Constr.*, vol. 67, pp. 1–21, Jul. 2016.

[206] M. Bremer, A. Mayr, V. Wichmann, K. Schmidtner, and M. Rutzinger, "A new multi-scale 3D-GIS-approach for the assessment and dissemination of solar income of digital city models," *Comput. Environ. Urban Syst.*, vol. 57, pp. 144–154, May 2016.

[207] Y. Song *et al.*, "Trends and Opportunities of BIM-GIS Integration in the Architecture, Engineering and Construction Industry: A Review from a Spatio-Temporal Statistical Perspective," *ISPRS Int. J. Geo-Information*, vol. 6, no. 12, p. 397, Dec. 2017.

[208] F. Biljecki, H. Ledoux, J. Stoter, and G. Vosselman, "The variants of an LOD of a 3D

building model and their influence on spatial analyses," *ISPRS J. Photogramm. Remote Sens.*, vol. 116, pp. 42–54, Jun. 2016.

[209] T. A. Nguyen-Gia, M. S. Dao, and C. Mai-Van, "A comparative survey of 3D GIS models," *2017 4th NAFOSTED Conf. Inf. Comput. Sci. NICS 2017 - Proc.*, vol. 2017–Janua, no. October, pp. 126–131, 2017.

[210] G. Landeschi, N. Dell'Unto, K. Lundqvist, D. Ferdani, D. M. Campanaro, and A. M. Leander Touati, "3D-GIS as a platform for visual analysis: Investigating a Pompeian house," *J. Archaeol. Sci.*, vol. 65, pp. 103–113, Jan. 2016.

[211] K. Iwamura, K. Muro, N. Ishimaru, and M. Fukushima, "4D-GIS (4 dimensional GIS) as spatial-temporal data mining platform and its application to managementand monitoring of large-scale infrastructures," *ICSDM 2011 - Proc. 2011 IEEE Int. Conf. Spat. Data Min. Geogr. Knowl. Serv.*, pp. 38–43, 2011.

[212] L. Lopez-Gonzalez, R. Otero de Cosca, M. Gomez-Heras, and S. Garcia-Morales, "A 4D GIS methodology to study variations in evaporation points on a heritage building," *Environ. Earth Sci.*, vol. 75, no. 14, 2016.

[213] N. G. Smith, M. Howland, and T. E. Levy, "Digital archaeology field recording in the 4th dimension: ArchField C++ a 4D GIS for digital field work," *2015 Digit. Herit. Int. Congr. Digit. Herit. 2015*, pp. 251–258, 2015.

[214] K. Iwamura, "Implementations of Spatio-Temporal Data Structure for Geographic Information System."

[215] Gopal M. Naik, Aditya M., and Suma B. Naik, "GIS Based 4D Model Development for Planning and Scheduling of a Construction Project ," *Int. J. Innov. Manag. Technol.*, vol. 2, no. 6, pp. 447–451, 2011.

[216] A. C. Kumar and T. Reshma, "4D Applications of GIS in Construction Management," *Adv. Civ. Eng.*, vol. 2017, pp. 1–9, 2017.

[217] R. Ghose and T. Welcenbach, "'Power to the people': Contesting urban poverty and power inequities through open GIS," *Can. Geogr.*, vol. 62, no. 1, pp. 67–80, 2018.

[218] Open Geospatial Consortium, "Welcome to the OGC," .

[219] V. Maquil, U. Leopold, L. M. De Sousa, L. Schwartz, and E. Tobias, "Towards a framework for geospatial tangible user interfaces in collaborative urban planning," *J. Geogr. Syst.*, vol. 20, no. 2, pp. 185–206, 2018.

[220] L. Hu *et al.*, "Developing geospatial Web service and system for SMAP soil moisture monitoring," in *2017 6th International Conference on Agro-Geoinformatics, Agro-Geoinformatics 2017*, 2017, pp. 1–5.

[221] I. Prieto, J. L. Izkara, and R. Béjar, "Web-based tool for the sustainable refurbishment in historic districts based on 3D city model," in *Lecture Notes in Geoinformation and Cartography*, no. 9783319256894, 2017, pp. 159–169.

[222] K. Schröter *et al.*, "Flood loss estimation using 3D city models and remote sensing data," *Environ. Model. Softw.*, vol. 105, pp. 118–131, Jul. 2018.

[223] M. S. B. PhridviRaj and C. V. GuruRao, "Data Mining – Past, Present and Future – A

Typical Survey on Data Streams," *Procedia Technol.*, vol. 12, pp. 255–263, 2014.

[224] D. J. Hand, "Principles of data mining," in *Drug Safety*, 2007, vol. 30, no. 7, pp. 621–622.

[225] X. Wu, X. Zhu, G.-Q. Wu, and W. Ding, "Data Mining with Big Data," *Knowl. Data Eng. IEEE Trans.*, vol. 26, no. 1, pp. 97–107, 2014.

[226] S. Sayad, "Data Mining Map," 2014. [Online]. Available: http://chem-eng.utoronto.ca/~datamining/dmc/data_mining_map.htm. [Accessed: 31-May-2018].

[227] C. Aggarwal, *Data Mining: The Textbook*, vol. 53, no. 9. 2013.

[228] S. S. Gulluoglu, "Segmenting customers with data mining techniques," in *2015 Third International Conference on Digital Information, Networking, and Wireless Communications (DINWC)*, 2015, pp. 154–159.

[229] L. Wang, Z. Fan, X. Wang, and L. Yang, "Text mining-based evaluation of the user experience in online shopping for clothing," *1st Int. Conf. Electron. Instrum. Inf. Syst. EIIS 2017*, vol. 2018–Janua, pp. 1–4, 2018.

[230] Z. Xiang, Z. Schwartz, J. H. Gerdes, and M. Uysal, "What can big data and text analytics tell us about hotel guest experience and satisfaction?," *Int. J. Hosp. Manag.*, vol. 44, pp. 120–130, Jan. 2015.

[231] J. Howe, "The Rise of Crowdsourcing," *Wired Mag.*, vol. 14, no. 06, pp. 1–5, 2006.

[232] W. Fan, Q. Li, J. Han, B. Zhao, and J. Gao, "Mining Reliable Information from Passively and Actively Crowdsourced Data," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16*, 2016, pp. 2121–2122.

[233] J. Surowiecki, "The Wisdom of Crowds," *How Collect. Wisdom Shapes Bus. Econ. Soc. Nations New York Doubleday*, p. 296, 2004.

[234] G. Le Bon, *The Crowd: A Study of the Popular Mind*. 1896.

[235] C. Mackay, *Extraordinary popular delusions and the madness of crowds.* 2015.

[236] G. Xintong, W. Hongzhi, Y. Song, and G. Hong, "Brief survey of crowdsourcing for data mining," *Expert Syst. Appl.*, vol. 41, no. 17, pp. 7987–7994, Dec. 2014.

[237] K. Domdouzis, B. Akhgar, S. Andrews, H. Gibson, and L. Hirsch, "A social media and crowdsourcing data mining system for crime prevention during and post-crisis situations," *J. Syst. Inf. Technol.*, vol. 18, no. 4, pp. 364–382, 2016.

[238] M. Erraguntla, D. Delen, R. K. Agrawal, K. Madanagopal, and R. Mayer, "Mobile-Based Sidewalk Inventory App for Smart Communities, Health, and Safety," *Suburb. Sustain.*, vol. 5, no. 1, 2017.

[239] M. Nik-Bakht and T. E. El-Diraby, "Sus-tweet-ability: Exposing public community's perspective on sustainability of urban infrastructure through online social media," *Int. J. Hum. Comput. Stud.*, vol. 89, pp. 54–72, May 2016.

[240] A. Katal, M. Wazid, and R. H. Goudar, "Big data: Issues, challenges, tools and Good

practices," *2013 6th Int. Conf. Contemp. Comput. IC3 2013*, pp. 404–409, 2013.

[241] B. Delibašić *et al.*, "Decision Support Systems V - Big Data Analytics for Decision Making: First International Conference, ICDSST 2015 Belgrade, Serbia, May 27-29, 2015 Proceedings," *Lect. Notes Bus. Inf. Process.*, vol. 216, pp. 10–21, 2015.

[242] D. Reinsel, J. Gantz, and J. Rydning, "Data Age 2025 : Don 't Focus on Big Data; Focus on the Data That's Big Data Age 2025 :," *IDC White Pap. Spons. by Seagate*, no. April, pp. 1–25, 2017.

[243] Economist Intelligence Unit, "The Deciding Factor: Big Data & Decision Making," *Econ.*, pp. 1–24, 2012.

[244] H. Zhang and L. Zhu, "Internet of Things: Key technology, architecture and challenging problems," *Proc. - 2011 IEEE Int. Conf. Comput. Sci. Autom. Eng. CSAE 2011*, vol. 4, pp. 507–512, 2011.

[245] R. H. Weber, "Internet of Things - New security and privacy challenges," *Comput. Law Secur. Rev.*, vol. 26, no. 1, pp. 23–30, 2010.

[246] R. Jalali, K. El-Khatib, and C. McGregor, "Smart city architecture for community level services through the internet of things," *2015 18th Int. Conf. Intell. Next Gener. Networks, ICIN 2015*, pp. 108–113, 2015.

[247] M. Milenkovic, "A Case for Interoperable IoT Sensor Data and Meta-data Formats," *Ubiquity*, vol. 2015, no. November, pp. 1–7, Nov. 2015.

[248] J. Manyika *et al.*, "McKinsey Global Institute: Unlocking the potential of the Internet of Things," *Digital McKinsey*, 2015. [Online]. Available: http://www.mckinsey.com/business-functions/business-technology/our-insights/the-internet-of-things-the-value-of-digitizing-the-physical-world. [Accessed: 12-Oct-2018].

[249] Y. Rezgui, C. J. Hopfe, and C. Vorakulpipat, "Generations of knowledge management in the architecture, engineering and construction industry: An evolutionary perspective," *Adv. Eng. Informatics*, vol. 24, no. 2, pp. 219–228, Apr. 2010.

[250] N. Dawood and S. Sikka, "Engineering, Construction and Architectural Management Development of 4D based performance indicators in construction industry," *Benchmarking An Int. J.*, vol. 16, no. 2, pp. 438–458, 2009.

[251] PlanGrid, "Construction Disconnected," 2018.

[252] J. Teizer and P. A. Vela, "Personnel tracking on construction sites using video cameras," *Adv. Eng. Informatics*, vol. 23, no. 4, pp. 452–462, Oct. 2009.

[253] W. Shen *et al.*, "Systems integration and collaboration in architecture, engineering, construction, and facilities management: A review," *Adv. Eng. Informatics*, vol. 24, no. 2, pp. 196–207, Apr. 2010.

[254] J. Irizarry, E. P. Karan, and F. Jalaei, "Integrating BIM and GIS to improve the visual monitoring of construction supply chain management," *Autom. Constr.*, vol. 31, pp. 241–254, 2013.

[255] D. A. Young, C. T. Haas, P. Goodrum, and C. Caldas, "Improving Construction Supply

Network Visibility by Using Automated Materials Locating and Tracking Technology," *J. Constr. Eng. Manag.*, vol. 137, no. 11, pp. 976–984, Nov. 2011.

[256] S. Hou, "An ontology-based holistic approach for multi-objective sustainable structural design," 2015.

[257] M. P. Gallaher, A. C. O'Connor, J. L. Dettbarn, Jr., and L. T. Gilday, "Cost Analysis of Inadequate Interoperability in the U.S. Capital Facilities Industry," Aug. 2004.

[258] McGraw-Hill Construction, "The Business Value of BIM in Europe SmartMarket Report," 2010.

[259] M. F. Muller, A. Garbers, F. Esmanioto, N. Huber, E. R. Loures, and O. Canciglieri, "Data interoperability assessment though IFC for BIM in structural design – a five-year gap analysis," *J. Civ. Eng. Manag.*, vol. 23, no. 7, pp. 943–954, Oct. 2017.

[260] J. Park, H. Cai, P. S. Dunston, and H. Ghasemkhani, "Database-Supported and Web-Based Visualization for Daily 4D BIM," *J. Constr. Eng. Manag.*, vol. 143, no. 10, p. 04017078, Oct. 2017.

[261] N. Hudson *et al.*, "Closed BIM Versus openBIM. The Opportunities, Challenges and Implications of openBIM," *BIM Journal*, 2017.

[262] Building SMART, "Open Standards - the basics," 2018. [Online]. Available: https://www.buildingsmart.org/standards/technical-vision/open-standards/. [Accessed: 17-May-2018].

[263] M. Das, J. C. Cheng, and S. S. Kumar, "Social BIMCloud: a distributed cloud-based BIM platform for object-based lifecycle information exchange," *Vis. Eng.*, vol. 3, no. 1, p. 8, Dec. 2015.

[264] S. Steiniger and A. J. S. Hunter, "The 2012 free and open source GIS software map - A guide to facilitate research, development, and adoption," *Comput. Environ. Urban Syst.*, vol. 39, pp. 136–150, 2013.

[265] S.-C. Yu and T.-A. Teo, "The Generalization of Bim/Ifc Model for Multi-Scale 3D Gis/Citygml Models," in *Asian Association on remote sensing*, 2014.

[266] A.-H. Hor, "A semantic Web platform for BIM-GIS integration," no. April 2015, 2016.

[267] E. P. Karan and J. Irizarry, "Extending BIM interoperability to preconstruction operations using geospatial analyses and semantic web services," *Autom. Constr.*, vol. 53, pp. 1–12, May 2015.

[268] A. Rafiee, E. Dias, S. Fruijtier, and H. Scholten, "From BIM to Geo-analysis: View Coverage and Shadow Analysis by BIM/GIS Integration," *Procedia Environ. Sci.*, vol. 22, no. 22, pp. 397–402, 2014.

[269] S. Tang, D. R. Shelden, C. M. Eastman, P. Pishdad-Bozorgi, and X. Gao, "A review of building information modeling (BIM) and the internet of things (IoT) devices integration: Present status and future trends," *Autom. Constr.*, vol. 101, pp. 127–139, May 2019.

[270] B. Dave, A. Buda, A. Nurminen, and K. Främling, "A framework for integrating BIM and IoT through open standards," *Autom. Constr.*, vol. 95, pp. 35–45, Nov. 2018.

[271] T. Berners-Lee, “Information Management: A Proposal,” *Word J. Int. Linguist. Assoc.*, vol. February 2, no. May, pp. 1–10, 1989.

[272] B. Marr, “Big Data: 20 Mind-Boggling Facts Everyone Must Read,” *www.forbes.com*, 2015. [Online]. Available: https://www.forbes.com/sites/bernardmarr/2015/09/30/big-data-20-mind-boggling-facts-everyone-must-read/#78b45a6c17b1. [Accessed: 08-Oct-2018].

[273] V. Nedeva and S. Dineva, “Intelligent e-Learning with New Web Technologies,” *Proc. 10Th Int. Conf. Virtual Learn.*, no. September, pp. 68–74, 2015.

[274] C. Bizer and T. Heath, “Linked Data -The Story So Far,” *Int. J. Semant. Web Inf. Syst.*, 2009.

[275] DBpedia, “DBpedia,” 2018. [Online]. Available: https://wiki.dbpedia.org/. [Accessed: 09-Oct-2018].

[276] L. Obrst, H. Liu, R. Wray, and L. Wilson, “Ontologies for semantically interoperable electronic commerce,” *IFIP Adv. Inf. Commun. Technol.*, vol. 108, pp. 325–333, 2003.

[277] A. Ouskel and A. Sheth, “Semantic Interoperability in Global Information Systems,” *SIGMOD Rec.*, vol. 28, no. 1, pp. 5–12, 1999.

[278] T. Barners-Lee, N. Shadbolt, and W. Hall, “The Semantic Web Revisited,” *IEEE Intelligent Systems*, vol. 21, no. 3. pp. 96–101, 2006.

[279] A. Maedche and P. Weiß, “Towards Ontology-Based Smart Organizations,” in *IFIP International Federation for Information Processing 2002*, 2013, vol. 85, pp. 201–208.

[280] R. J. Rabelo, S. Gusmeroli, C. Arana, and T. Nagellen, “The Ecolead ICT Infrastructure For Collaborative Networked Organizations,” *Network-Centric Collab. Support. Fram.*, vol. 224, no. 3, pp. 451–460, 2007.

[281] W. Mulder, P. H. H. Rongen, and G. R. Meijer, “Towards Ontology-Based CNO Matching Applied to Squads,” *Collab. Networks Their Breed. Environ.*, vol. 186, pp. 117–124, 2006.

[282] C. Kuster, J.-L. Hippolyte, and Y. Rezgui, “Collaborative Network for District Energy Operation and Semantic Technologies: A Case Study,” in *Camarinha-Matos L., Afsarmanesh H., Rezgui Y. (eds) Collaborative Networks of Cognitive Systems. PRO-VE 2018.*, 2018, pp. 486–495.

[283] L. M. Camarinha-Matos and H. Afsarmanesh, “Collaborative networks: A new scientific discipline,” *Virtual Organ. Syst. Pract.*, vol. 16, no. 4–5, pp. 73–80, Oct. 2005.

[284] National Research Council, “National Research Council. (1998). Visionary Manufacturing Challenges for 2020, Visionary manufacturing challenges for 2020/Committee on Visionary Manufacturing Challenges, Board on Manufacturing and Engineering Design, Commission on Engineering and Techn,” in *Board on manufacturing and engineering design commission on engineering and technical systems*, 1998.

[285] L. M. Camarinha-Matos, H. Afsarmanesh, N. Galeano, and A. Molina, “Collaborative

networked organizations - Concepts and practice in manufacturing enterprises," *Comput. Ind. Eng.*, vol. 57, no. 1, pp. 46–60, Aug. 2009.

[286] N. Silva and J. Rocha, "VE Infrastructures Requirements for Cooperation and Knowledge Sharing," in *IFIP International Federation for Information Processing*, 2013, vol. 56, pp. 79–86.

[287] M. Keller, K. Menzel, and R. J. Scherer, "Towards a Meta-Model for Collaborative Construction Project Management," *Collab. Networks Their Breed. Environ.*, vol. 186, pp. 361–368, 2006.

[288] L. Allan and K. Menzel, "Virtual enterprises for integrated energy service provision," *IFIP Adv. Inf. Commun. Technol.*, vol. 307, pp. 659–666, 2009.

[289] I. Praça, H. Morais, M. Cardoso, C. Ramos, and Z. Vale, "Virtual power producers integration into MASCEM," in *IFIP International Federation for Information Processing*, vol. 243, Boston, MA: Springer US, 2007, pp. 291–298.

[290] USEF Foundation, *USEF The Framework Explained A solid foundation for smart energy futures*. 2015.

[291] F. Rahimi and A. Ipakchi, "Demand response as a market resource under the smart grid paradigm," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 82–88, 2010.

[292] V. Guevara-Masis, H. Afsarmanesh, and L. O. Hertzberger, "Ontology-Based Automatic Data Structure Generation for Collaborative Networks," in *PRO-VE 2004:Virtual Enterprises and Collaborative Networks*, 2004, pp. 163–174.

[293] G. Anzaldi *et al.*, "Knoholem: Knowledge-Based Energy Management for Public Buildings Through Holistic Information Modeling and 3D Visualization," vol. 70.

[294] W3C, "Semantic Web - W3C," 2018. [Online]. Available: https://www.w3.org/standards/semanticweb/. [Accessed: 09-Oct-2018].

[295] T. R. Gruber, "A translation approach to portable ontology specifications," *Knowl. Acquis.*, vol. 5, no. 2, pp. 199–220, 1993.

[296] M. Uschold and M. Gruninger, "Ontologies and semantics for seamless connectivity," *ACM SIGMOD Rec.*, vol. 33, no. 4, p. 58, 2005.

[297] T. Berners and R. Fielding, "Uniform Resource Identifier (URI): Generic Syntax Status," 2013.

[298] W3C/IETF, "URIs, URLs, and URNs: clarifications and recommendations 1.0," 2001.

[299] W. F. Hammond, "Extensible Markup Language (XML) Standard Generalized Markup Language (SGML)."

[300] G. Schreiber and Y. Raimond, "Rdf 1.1 Primer," *W3C Working Group Note 24 June 2014*, 2014. [Online]. Available: https://www.w3.org/TR/rdf11-primer/. [Accessed: 10-Oct-2018].

[301] W3C, "RDF Schema 1.1," 2014.

[302] M. K. Smith, C. Welty, and D. L. McGuinness, "OWL web ontology language guide. W3C Recommendation, 10 February 2004," *World Wide Web Consortium*. 2004.

[303] G. Pandey, “The Semantic Web: An Introduction and Issues,” *Int. J. Eng. Res. Appl.*, vol. 2, no. 1, pp. 780–786, 2012.

[304] F. Baader, “What’s new in Description Logics,” *Informatik-Spektrum*, vol. 34, no. 5, pp. 434–442, 2011.

[305] W3C Group, “OWL 2 Web Ontology Language Document Overview,” *W3C Recommendation 11 December 2012*, 2012. [Online]. Available: https://www.w3.org/TR/owl2-overview/. [Accessed: 11-Oct-2018].

[306] B. Motik, B. C. Grau, I. Horrocks, Z. Wu, A. Fokoue, and C. Lutz, “OWL 2 web ontology language: Profiles - W3C Recommendation,” 2012.

[307] B. Glimm and C. Ogbuji, “SPARQL 1.1 Entailment Regimes W3C Recommendation 21 March 2013,” *W3C*. Addison-Wesley, p. 1462, 2013.

[308] B. Glimm and C. Ogbuji, “SPARQL 1.1 Entailment Regimes. Technical report,” 2012.

[309] M. Horridge and M. Musen, “Snap-SPARQL: A java framework for working with SPARQL and OWL,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9557, pp. 154–165, 2016.

[310] Jennifer Zaino, “Semantic Web and Semantic Technology Trends in 2018,” *Dataversity*, 2017. [Online]. Available: http://www.dataversity.net/semantic-technology-semantic-web-trends-2018/. [Accessed: 11-Oct-2018].

[311] B. Hassan, “Towards Semantic Web: Challenges and Needs,” *Int. J. Eng. Comput. Sci.*, no. April, pp. 4–8, 2015.

[312] K. Aberer, M. Hauswirth, and A. Salehi, “Middleware support for the ‘Internet of Things,’” *5th GI/ITG KuVS Fachgespräch “Drahtlose Sensornetze,”* no. 5005, p. 5, 2006.

[313] J. P. Calbimonte, S. Sarni, J. Eberle, and K. Aberer, “XGSN: An open-source semantic sensing middleware for the web of things,” *CEUR Workshop Proc.*, vol. 1401, pp. 51–66, 2014.

[314] L. Battle *et al.*, “Building the Internet of Things Using RFID: The RFID Ecosystem Experience,” *IEEE Internet Comput.*, vol. 13, no. 3, pp. 48–55, May 2009.

[315] V.-M. Scuturici, S. Surdu, Y. Gripay, and J.-M. Petit, “UbiWare : Web-based Dynamic Data & Service Management Platform for AmI,” in *Proceedings of the Posters and Demo Track on - Middleware ’12*, 2012, pp. 5–6.

[316] L. Roalter, L. Roalter, M. Kranz, A. Möller, and T. U. München, “A middleware for intelligent environments and the internet of things,” *Ubiquitous Intell. Comput.*, pp. 267–281, 2010.

[317] M. Quigley *et al.*, “ROS: an open-source Robot Operating System,” in *ICRA Workshop on Open Source Software*.

[318] S. Liang, C. Y. Huang, and T. Khalafbeigi, “OGC SensorThings API Part 1: Sensing,” *Open Geospatial Consortium. Implement. Stand.*, pp. 1–105, 2016.

[319] M. Botts and C. A. Robin, “Open Geospatial Consortium OGC ® SensorML: Model and XML Encoding Standard,” p. 192, 2014.

[320] OGC, "OGC SWE Common Data Model Encoding Standard," *OpenGIS® Implement. Stand.*, 2011.

[321] Open Geospatial Consortium, "ISO 19156: 2011-Geographic information: Observations and measurements," vol. 2011, p. 54, 2011.

[322] W3C, "Semantic Sensor Network Ontology," 2005. [Online]. Available: https://www.w3.org/2005/Incubator/ssn/ssnx/ssn%5Cnhttp://www.w3.org/2005/Incubator/ssn/ssnx/ssn.

[323] M. Compton *et al.*, "The SSN ontology of the W3C semantic sensor network incubator group," *J. Web Semant.*, vol. 17, pp. 25–32, 2012.

[324] University of Trento, "DOLCE: Descriptive Ontology for Linguistic and Cognitive Engineering," 2004. [Online]. Available: http://www.loa.istc.cnr.it/old/DOLCE.html. [Accessed: 13-Oct-2018].

[325] V. Mascardi, V. Cordì, and P. Rosso, "A Comparison of Upper Ontologies (Technical Report DISI-TR-06-21)," *Dip. di Inform. e Sci. dell'Informazione*, pp. 55–64, 2007.

[326] A. Haller, K. Janowicz, S. Cox, D. Le Phuoc, K. Taylor, and M. Lefrançois, "Semantic Sensor Network Ontology." W3C. 19 October 2017., 2017.

[327] S. Cox, C. Little, J. R. Hobbs, and F. Pan, "Time Ontology in OWL. W3C Recommendation." W3C. 19 October 2017, 2017.

[328] L. Daniele, "The Smart Appliances REFerence (SAREF) ontology Documentation," 2015. [Online]. Available: https://w3id.org/saref. [Accessed: 13-Oct-2018].

[329] ETSI, "TS 103 264 - V1.1.1 - SmartM2M; Smart Appliances; Reference Ontology and oneM2M Mapping," vol. 1, pp. 1–26, 2015.

[330] L. Daniele, F. den Hartog, and J. Roes, "Created in Close Interaction with the Industry: The Smart Appliances REFerence (SAREF) Ontology," in *Lecture Notes in Business Information Processing*, vol. 225, 2015, pp. 100–112.

[331] E. Kharlamov *et al.*, "Semantic access to streaming and static data at Siemens," *J. Web Semant.*, vol. 44, pp. 54–74, 2017.

[332] G. Atemezing *et al.*, "Transforming meteorological data into linked data," *Semant. Web*, vol. 4, no. 3, pp. 285–290, 2013.

[333] International Organization for Standardization (ISO), *ISO 16739-1:2018 - Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries*. 2018.

[334] S. Mihindu and Y. Arayici, "Digital construction through BIM systems will drive the Re-engineering of construction business practices," *Proc. - Int. Conf. Vis. VIS 2008, Vis. Built Rural Environ.*, pp. 29–34, 2008.

[335] P. Pauwels, S. Zhang, and Y. C. Lee, "Semantic web technologies in AEC industry: A literature overview," *Autom. Constr.*, vol. 73, pp. 145–165, Jan. 2017.

[336] C. J. Anumba, J. Pan, R. R. A. Issa, and I. Mutis, "Collaborative project information management in a semantic web environment," *Eng. Constr. Archit. Manag.*, vol. 15, no. 1, pp. 78–94, Jan. 2008.

[337] P. Pauwels and W. Terkaj, “EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology,” *Autom. Constr.*, vol. 63, pp. 100–133, Mar. 2016.

[338] N. V. Hoang and S. Törmä, “Implementation and Experiments with an IFC-to-Linked Data Converter.”

[339] P. Pauwels and D. Van Deursen, “First International Workshop on Linked Data in Architecture and Construction-Workshop Report,” in *LDAC2012*, 2012.

[340] R. Barbau *et al.*, “OntoSTEP: Enriching product model data using ontologies,” *CAD Comput. Aided Des.*, vol. 44, no. 6, pp. 575–590, 2012.

[341] N.V. Hoang, “IFC-to-linked data conversion: multilayering approach,” in *LDAC 2015 3rd International Workshop on Linked Data in Architecture and Construction (LDAC)*, 2015.

[342] R. Y. M. Program and M. Info, “Opening BIM to the Web – IFC-to-RDF Conversion Software,” 2013.

[343] Pieter Pauwels, “IFC-RDF convertor repository.” [Online]. Available: http://smartlab1.elis.ugent.be:8889/IFC-repo/. [Accessed: 09-May-2018].

[344] P. Pauwels, T. Krijnen, W. Terkaj, and J. Beetz, “Enhancing the ifcOWL ontology with an alternative representation for geometric data,” *Autom. Constr.*, vol. 80, pp. 77–94, Aug. 2017.

[345] R. Howard and B. C. Björk, “Building information modelling - Experts’ views on standardisation and industry deployment,” *Adv. Eng. Informatics*, vol. 22, no. 2, pp. 271–280, Apr. 2008.

[346] Y. Rezgui, T. Beach, and O. Rana, “a Governance Approach for Bim Management Across Lifecycle and Supply Chains Using Mixed-Modes of Information Delivery,” *J. Civ. Eng. Manag.*, vol. 19, no. 2, pp. 239–258, Apr. 2013.

[347] H. Ashcraft, “Building information modeling: A framework for collaboration,” *Constr. Lawyer*, vol. 28, no. 3, pp. 1–14, 2008.

[348] MHC, “The Business Value of BIM for Construction in Major Global Markets,” 2014.

[349] X. Li and A. G.-O. Yeh, “Zoning land for agricultural protection by the integration of remote sensing, GIS, and cellular automata,” *Photogrammetric Engineering and Remote Sensing*, vol. 67, no. 4. pp. 471–477, 2001.

[350] J. Khan and M. H. Arsalan, “Estimation of rooftop solar photovoltaic potential using geo-spatial techniques: A perspective from planned neighborhood of Karachi - Pakistan,” *Renew. Energy*, vol. 90, pp. 188–203, 2016.

[351] X. Mo, S. Liu, Z. Lin, Y. Xu, Y. Xiang, and T. R. McVicar, “Prediction of crop yield, water consumption and water use efficiency with a SVAT-crop growth model using remotely sensed data on the North China Plain,” *Ecol. Modell.*, vol. 183, no. 2–3, pp. 301–322, 2005.

[352] K. Sanga-Ngoie, K. Iizuka, and S. Kobayashi, “Estimating CO2 sequestration by forests in oita prefecture, Japan, by combining LANDSAT ETM+ and ALOS satellite remote

sensing data," *Remote Sens.*, vol. 4, no. 11, pp. 3544–3570, 2012.

[353] M. Perry and J. Herring, "OGC GeoSPARQL-A geographic query language for RDF data," *OGC Candidate Implement. Stand.*, p. 57, 2012.

[354] R. Battle and D. Kolas, "Linking geospatial data with GeoSPARQL," *Semant. Web*, vol. 0, no. 0, pp. 1–11, 2011.

[355] GeoNames, "GeoNames," 2018. [Online]. Available: http://www.geonames.org/. [Accessed: 09-Oct-2018].

[356] K. Su, J. Li, and H. Fu, "Smart city and the applications," in *2011 International Conference on Electronics, Communications and Control (ICECC)*, 2011, pp. 1028–1031.

[357] CityGML, "CityGML." [Online]. Available: https://www.citygml.org/. [Accessed: 26-Feb-2018].

[358] E. Standard, T. H. Kolbe, C. Nagel, and E. Standard, "Open Geospatial Consortium OGC City Geography Markup Language (CityGML) En- coding Standard," 2012.

[359] W. Y. Lin, P. H. Lin, and H. P. Tserng, "Automating the Generation of Indoor Space Topology for 3D Route Planning Using BIM and 3D-GIS Techniques," in *Proceedings of the 34th International Symposium on Automation and Robotics in Construction (ISARC)*, 2017.

[360] C. Metral, G. Falquet, and A. F. Cutting-Decelle, "Towards semantically enriched 3d city models : an ontology-based approach," *Proceeding GeoWeb 2009 Acad. Track – Cityscapes - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, pp. 40–45, 2009.

[361] O. Zalamea *et al.*, "From a CityGML to an ontology-based approach to support preventive conservation of built cultural heritage," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7616, no. April 2016, pp. 73–90, 2013.

[362] M. Poveda-Villalón, R. García-Castro, and A. Gómez-Pérez, "Building an ontology catalogue for smart cities," in *eWork and eBusiness in Architecture, Engineering and Construction*, 2014, pp. 829–836.

[363] Ontology Engineering Group, "Ontology Catalogue for Smart Cities," 2018. [Online]. Available: http://smartcity.linkeddata.es/datasets/index.html. [Accessed: 15-Oct-2018].

[364] M. S. Fox, "The PolisGnosis Project Enabling the Computational Analysis of City Performance."

[365] International Standard Organization, "ISO 37120 Sustainable development of communities — Indicators for city services and quality of life," vol. 37120. p. 14, 2014.

[366] A. Psyllidis, "Ontology-Based Data Integration from Heterogeneous Urban Systems: A Knowledge Representation Framework for Smart Cities," *14th Int. Conf. Comput. Urban Plan. Urban Manag. (CUPUM 2015)*, 2015.

[367] V. Corrado, I. Ballarini, L. Madrazo, and G. Nemirovskij, "Data structuring for the

ontological modelling of urban energy systems: The experience of the SEMANCO project," *Sustain. Cities Soc.*, vol. 14, no. 1, pp. 223–235, 2015.

[368] J. L. Hippolyte, Y. Rezgui, H. Li, B. Jayan, and S. Howell, "Ontology-driven development of web services to support district energy applications," *Autom. Constr.*, vol. 86, pp. 210–225, 2018.

[369] A. P. Bellini, P. Nesi, and N. Rauch, "Km4City - the Knowledge Model 4 the City Smart City Ontology," pp. 1–28, 2015.

[370] D. Corsar, M. Markovic, P. Edwards, and J. D. Nelson, "The Transport Disruption Ontology."

[371] R. Hodgson and P. J. Keller, "QUDT-quantities, units, dimensions and data types in OWL and XML," *Online (September 2011) http://www.qudt.org*, p. 34, 2011.

[372] E. Simperl, "Reusing ontologies on the Semantic Web: A feasibility study," *Data Knowl. Eng.*, vol. 68, no. 10, pp. 905–925, 2009.

[373] M. Uschold, "Building Ontologies : Towards a United Methodology," in *Expert Systems 96*, 1996, no. September.

[374] M. Uschold and M. King, "Towards a Methodology for Building Ontologies," 1995.

[375] M. Uschold, M. Healy, K. Williamson, P. Clark, and S. Woods, "Ontology reuse and application," *Form. Ontol. Inf. Syst. IOS Press*, vol. 179, no. January 2000, pp. 179–194, 1998.

[376] M. Uschold and M. Gruninger, "Ontologies: principles, methods and applications," *Knowl. Eng. Rev.*, vol. 11, no. 02, p. 93, 2009.

[377] M. Fernandez, A. Gomez-Perez, and N. Juristo, "METHONTOLOGY: From Ontological Art Towards Ontological Engineering," in *Fourteenth National Conference on Artificial Intelligence, American Association for Artificial Intelligence*, 1997.

[378] M. d'Aquin, A. Gangemi, E. Motta, and M. Dzbor, "NeOn Tool Support for Building Ontologies by Reuse," *Framework*, pp. 3–4, 2009.

[379] M. C. Suárez-Figueroa, A. Gómez-Pérez, and M. Fernández-López, "The NeOn Methodology framework: A scenario-based methodology for ontology development," *Appl. Ontol.*, vol. 10, no. 2, pp. 107–145, Sep. 2015.

[380] M. C. Suárez-Figueroa, A. Gómez-Pérez, and M. Fernández-López, "The NeOn Methodology for Ontology Engineering," in *Ontology Engineering in a Networked World*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 9–34.

[381] M. del C. Suarez, "NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse," Universidad Politécnica de Madrid, 2010.

[382] M. Westerlund, S. Leminen, and M. Rajahonka, *Designing Business Models for the Internet of Things*, vol. 4, no. 7. Berlin, Heidelberg: Springer Berlin Heidelberg, 2018.

[383] Benson Chan, "The Smart City Ecosystem Framework - A Model for Planning Smart Cities," *Iot for All*. [Online]. Available: https://www.iotforall.com/smart-city-ecosystem-framework-model-for-planning-smart-cities/. [Accessed: 19-Oct-2018].

[384] N. Taylor Buck and A. While, “Competitive urbanism and the limits to smart city innovation: The UK Future Cities initiative,” *Urban Stud.*, vol. 54, no. 2, pp. 501–519, 2017.

[385] S. Murray, “Empowering cities,” *Econ. Intell. Unit Ltd.*, 2016.

[386] Municipality of Copenhagen, “Cita Data Exchange - Lessons Learned from a public/private data collaboration,” 2018.

[387] ⍰ ⍰ ⍰ ⍰, “Smart Cities Data Sharing Framework,” 2018.

[388] B. Martinez, A. Angles, I. Vilajosana, X. Vilajosana, J. Llosa, and M. Domingo-Prieto, “Bootstrapping smart cities through a self-sustainable model based on big data flows,” *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 128–134, 2013.

[389] N. Walravens and P. Ballon, “Platform business models for smart cities: From control and value to governance and public value,” *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 72–79, 2013.

[390] V. C. Storey and I. Y. Song, “Big data technologies and Management: What conceptual modeling can do,” *Data Knowl. Eng.*, vol. 108, pp. 50–67, Mar. 2017.

[391] G. Bello-Orgaz, J. J. Jung, and D. Camacho, “Social big data: Recent achievements and new challenges,” *Inf. Fusion*, vol. 28, pp. 45–59, 2016.

[392] R. G. Hollands, “Critical interventions into the corporate smart city,” *Cambridge J. Reg. Econ. Soc.*, vol. 8, no. 1, pp. 61–77, 2015.

[393] M. Provoost, “From Welfare City to Neoliberal Utopia,” *Strelka Talks*, 2013. [Online]. Available: https://vimeo.com/64392842.

[394] M. Foucault, “Surveiller et punir: Naissance de la prison,” 1975.

[395] G. Orwell, *1984*. 1946.

[396] Morozov, *To Save Everything Click Here*. 2013.

[397] U. Haque, “What Is a City that It Would Be ‘Smart’?,” *City a Box*, vol. 34, p. Alexander, C., Ishikawa, S., Silverstein, M. (19, 2012.

[398] Bsi, “Pd 8101 : 2014 Smart cities – Guide to the role of the planning and development process,” 2014.

[399] A. Meijer and M. P. R. Bolívar, “Governing the smart city: a review of the literature on smart urban governance,” *Int. Rev. Adm. Sci.*, vol. 82, no. 2, pp. 392–408, 2016.

[400] G. V. Pereira, M. A. Macadar, E. M. Luciano, and M. G. Testa, “Delivering public value through open government data initiatives in a Smart City context,” *Inf. Syst. Front.*, vol. 19, no. 2, pp. 213–229, 2017.

[401] R. E. Sieber and P. A. Johnson, “Civic open data at a crossroads: Dominant models and current challenges,” *Gov. Inf. Q.*, vol. 32, no. 3, pp. 308–315, Jul. 2015.

[402] A. Gritzalis, A. Tsohou, and C. Lambrinoudakis, “Transparency-enabling systems for open governance: Their impact on citizens’ trust and the role of information privacy,” in *Communications in Computer and Information Science*, vol. 792, 2017,

pp. 47–63.

[403] K. Paskaleva, A. Karvonen, T. Linjordet, J. Evans, D. Yang, and C. Martin, "Data Governance in the Sustainable Smart City," *Informatics*, vol. 4, no. 4, p. 41, 2017.

[404] Sunlight Foundation, "How Unique is the New U.S. Open Data Policy? : Sunlight Foundation," 2013. [Online]. Available: https://sunlightfoundation.com/2013/05/13/how-unique-is-the-new-u-s-open-data-policy/. [Accessed: 30-Oct-2018].

[405] European Commission, "Open data: An engine for innovation, growth and transparent governance," 2011.

[406] Deloitte, "The challenge of paying for smart cities projects," 2018.

[407] S. Buckingham Shum *et al.*, "Towards a global participatory platform," *Eur. Phys. J. Spec. Top.*, vol. 214, no. 1, pp. 109–152, Nov. 2012.

[408] A. Goldstein, B., Dyson, L., & Nemani, *Beyond Transparency: Open Data and the Future of Civic Innovation*. 2013.

[409] Communities and Local Government and Department for Communities and Local Government, "A plain English guide to the Localism Bill," 2011.

[410] S. Hamilton and X. Zhu, "Funding and financing smart cities," 2018.

[411] Deloitte, "Forces of change: Smart cities," *Deloitte*, 2018.

[412] F. Gonzalez-Zapata and R. Heeks, "The multiple meanings of open government data: Understanding different stakeholders and their perspectives," *Gov. Inf. Q.*, vol. 32, no. 4, pp. 441–452, Oct. 2015.

[413] NYC Council, "NYC Participatory Budgeting," 2018. [Online]. Available: http://ideas.pbnyc.org/11/40.69678/-73.88306. [Accessed: 01-Jun-2018].

[414] City of Chicago, "311 Service Tracker," 2014. [Online]. Available: http://servicetracker.cityofchicago.org/. [Accessed: 01-Jun-2018].

[415] Sunlight Foundation, "Ten Principles for Opening Up Government Data." 2010.

[416] Las Vegas city council, "Las Vegas OpenData," 2018. [Online]. Available: https://opendata.lasvegasnevada.gov/browse. [Accessed: 01-Jun-2018].

[417] San Francisco City Council, "DataSF | San Francisco Open Data," 2018. [Online]. Available: https://datasf.org/opendata/. [Accessed: 01-Jun-2018].

[418] City of Edinburgh Council, "Datasets - City of Edinburgh Council," 2015. [Online]. Available: https://data.edinburghopendata.info//dataset. [Accessed: 01-Jun-2018].

[419] Open Knowledge Foundation, "US City Open Data Census," 2015. [Online]. Available: http://us-cities.survey.okfn.org/. [Accessed: 01-Jun-2018].

[420] J. Coplák and P. Raksanyi, *Planning sustainable settlements*. 2003.

[421] E. Turban, J. . Aronson, T.-P. Liang, and R. Sharda, *Decision Support and Business Intelligence Systems (8th Edition)*, vol. 8th. 2006.

[422] S. Alter, *Decision Support Systems: Current Practice and Continuing Challenges (Addison-Wesley series on decision support)*. 1979.

[423] D. Arnott, "A framework for developing intelligent decision support systems using evolutionary fuzzy cognitive maps.," *J. Intell. Fuzzy Syst.*, vol. 19, no. 2, pp. 151–170, 2008.

[424] E. De Kock, "Decentralising the Codification of Rules in a Decision Support Expert Knowledge Base," University of Pretoria, 2003.

[425] B. Peuportier, E. Popovici, and M. Trocmé, "Analyse de cycle de vie à l ' échelle du quartier - Bilan et perspectives du projet ADEQUA," *Ariadne*, pp. 1–8, 2006.

[426] G. Yepez, "CONSTRUCTION D'UN OUTIL D'EVALUATION ENVIRONNE- MENTALE DES ECOQUARTIERS : vers une méthode systé- mique de mise en oeuvre de la ville durable," 2011.

[427] Odysseus project, "Odysseus Framework Definition," 2013.

[428] A. Florea *et al.*, "Decision support tool for retrofitting a district towards district as a service," *Proc. - 2013 IEEE Int. Work. Intell. Energy Syst. IWIES 2013*, pp. 70–75, 2013.

[429] ECODISTR-ICT project, "Proposal for inclusion of KPIs , Graphical User-interface , and Technical Design of the Integrated Decision Support System," 2014.

[430] H. Rashed-Ali, "Spatial analysis for neighborhood sustainability assessment," *42nd ASES Natl. Sol. Conf. 2013, Sol. 2013, Incl. 42nd ASES Annu. Conf. 38th Natl. Passiv. Sol. Conf.*, pp. 107–114, 2013.

[431] L. Duclos, "Analyse de cycle de vie," 2011.

[432] Y. Diab *et al.*, "Analyse de cycle de vie à l'échelle du quartier : un outil d'aide à la décision? Le cas de la ZAC Claude Bernard à Paris (France)," *Environ. urbain*, vol. 5, p. c1, 2012.

[433] URB-Grade Project, "Decision Support Tool for Retrofitting a District, Towards the District as a Service D3.5.-PREDICTION MODULE SPECIFICATION AND DESIGN," pp. 1–29, 2015.

[434] A. Al-Ghandoor, J. O. Jaber, I. Al-Hinti, and I. M. Mansour, "Residential past and future energy consumption: Potential savings and environmental impact," *Renew. Sustain. Energy Rev.*, vol. 13, no. 6–7, pp. 1262–1274, 2009.

[435] A. D. Patton, A. F. Atiya, A. G. Parlos, J. Muthusami, and E. Oufi, "Development of an intelligent long-term electric load forecasting system," in *Journal of Dynamic Systems, Measurement, and Control*, 2002, vol. 131, no. 3, pp. 288–292.

[436] A. Azadeh and Z. S. Faiz, "A meta-heuristic framework for forecasting household electricity consumption," *Appl. Soft Comput. J.*, vol. 11, no. 1, pp. 614–620, 2011.

[437] Ü. B. Filik, Ö. N. Gerek, and M. Kurban, "A novel modeling approach for hourly forecasting of long-term electric energy demand," *Energy Convers. Manag.*, vol. 52, no. 1, pp. 199–211, 2011.

[438] J. C. Wang, "A study on the energy performance of school buildings in Taiwan,"

Energy Build., vol. 133, pp. 810–822, 2016.

[439] H. M. Al-Hamadi and S. A. Soliman, “Long-term/mid-term electric load forecasting based on short-term correlation and annual growth,” *Electr. Power Syst. Res.*, vol. 74, no. 3, pp. 353–361, 2005.

[440] S. Farzana, M. Liu, A. Baldwin, and M. U. Hossain, “Multi-model prediction and simulation of residential building energy in urban areas of Chongqing, South West China,” *Energy Build.*, vol. 81, pp. 161–169, 2014.

[441] L. Hernández *et al.*, “Artificial Neural Network for Short-Term Load Forecasting in Distribution Systems,” *Energies*, vol. 7, no. 3, pp. 1576–1598, 2014.

[442] R. Platon, V. R. Dehkordi, and J. Martel, “Hourly prediction of a building’s electricity consumption using case-based reasoning, artificial neural networks and principal component analysis,” *Energy Build.*, vol. 92, pp. 10–18, 2015.

[443] Y. H. Hsiao, “Household electricity demand forecast based on context information and user daily schedule analysis from meter data,” *IEEE Trans. Ind. Informatics*, vol. 11, no. 1, pp. 33–43, Feb. 2015.

[444] A. Garulli, S. Paoletti, and A. Vicino, “Models and Techniques for Electric Load Forecasting in the Presence of Demand Response,” *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 3, pp. 1087–1097, 2015.

[445] S. R. Twanabasu and B. A. Bremdal, “Load forecasting in a smart grid oriented building,” in *22nd International Conference and Exhibition on Electricity Distribution (CIRED 2013)*, 2013, pp. 0907–0907.

[446] R. X. Li, “Design and realization of 3-DOF welding manipulator control system based on motion controller,” *Energy Procedia*, vol. 14, pp. 931–936, 2012.

[447] J. Massana, C. Pous, L. Burgas, J. Melendez, and J. Colomer, “Short-term load forecasting in a non-residential building contrasting models and attributes,” *Energy Build.*, vol. 92, pp. 322–330, 2015.

[448] M. Mohandes, “Support vector machines for short-term electrical load forecasting,” *Int. J. Energy Res.*, vol. 26, no. 4, pp. 335–345, Mar. 2002.

[449] J. Y. Fan, J. D. Mcdonald, and S. Member, “A real-time implementation of short-term load forecasting for distribution power systems,” *IEEE Trans. Power Syst.*, vol. 9, no. 2, pp. 988–994, 1994.

[450] S. Gonzales Chavez, J. Xiberta Bernat, and H. Llaneza Coalla, “Forecasting of energy production and consumption in Asturias (northern Spain),” *Energy*, vol. 24, no. 3, pp. 183–198, 1999.

[451] E. A. Feinberg and D. Genethliou, “Peak demand control in commercial buildings with target peak adjustment based on load forecasting,” in *Applied Mathematics for Power Systems*, 2005, vol. 2, no. September, pp. 269–285.

[452] P. Chujai, N. Kerdprasop, and K. Kerdprasop, “Time series analysis of household electric consumption with ARIMA and ARMA models,” in *Lecture Notes in Engineering and Computer Science*, 2013, vol. 2202, pp. 295–300.

[453] S. Makridakis and M. Hibon, “The M3-Competition: results, conclusions and implications,” *Int. J. Forecast.*, vol. 16, no. 4, pp. 451–476, 2000.

[454] C. Kuster, Y. Rezgui, and M. Mourshed, “Electrical load forecasting models: A critical systematic review,” *Sustainable Cities and Society*, vol. 35. pp. 257–270, 2017.

[455] E. B. V. Elsevier, “About Scopus,” 2012. [Online]. Available: http://www.info.sciverse.com/scopus/about.

[456] R. E. Abdel-Aal, A. Z. Al-Garni, and Y. N. Al-Nassar, “Modelling and forecasting monthly electric energy consumption in eastern Saudi Arabia using abductive networks,” *Energy*, vol. 22, no. 9, pp. 911–921, 1997.

[457] M. Aydinalp, V. I. Ugursal, and A. S. Fung, “Modeling of the space and domestic hot-water heating energy-consumption in the residential sector using neural networks,” *Appl. Energy*, vol. 79, no. 2, pp. 159–178, 2004.

[458] M. Beccali, M. Cellura, V. Lo Brano, and A. Marvuglia, “Short-term prediction of household electricity consumption: Assessing weather sensitivity in a Mediterranean area,” *Renew. Sustain. Energy Rev.*, vol. 12, no. 8, pp. 2040–2065, 2008.

[459] V. Bianco, O. Manca, and S. Nardini, “Electricity consumption forecasting in Italy using linear regression models,” *Energy*, vol. 34, no. 9, pp. 1413–1421, 2009.

[460] F. Boulaire, A. Higgins, G. Foliente, and C. McNamara, “Statistical modelling of district-level residential electricity use in NSW, Australia,” *Sustain. Sci.*, vol. 9, no. 1, pp. 77–88, 2014.

[461] V. Cheng and K. Steemers, “Modelling domestic energy consumption at district scale: A tool to support national and local energy policies,” *Environ. Model. Softw.*, vol. 26, no. 10, pp. 1186–1198, 2011.

[462] L. Ciabattoni, M. Grisostomi, G. Ippoliti, and S. Longhi, “A Fuzzy Logic tool for household electrical consumption modeling,” in *IECON Proceedings (Industrial Electronics Conference)*, 2013, pp. 8022–8027.

[463] Z. Dilaver and L. C. Hunt, “Modelling and forecasting Turkish residential electricity demand,” *Energy Policy*, vol. 39, no. 6, pp. 3117–3127, 2011.

[464] H. Fan, I. F. MacGill, and A. B. Sproul, “Statistical analysis of driving factors of residential energy demand in the greater Sydney region, Australia,” *Energy Build.*, vol. 105, pp. 9–25, 2015.

[465] D. Fischer, A. Härtl, and B. Wille-Haussmann, “Model for electric load profiles with high time resolution for German households,” *Energy Build.*, vol. 92, pp. 170–179, 2015.

[466] M. Gul, S. A. Qazi, and W. A. Qureshi, “Incorporating economic and demographic variablesfor forecasting electricity consumption in Pakistan,” in *2011 2nd International Conference on Electric Power and Energy Conversion Systems, EPECS 2011*, 2011, pp. 1–5.

[467] R. Inglesi, “Aggregate electricity demand in South Africa: Conditional forecasts to 2030,” *Appl. Energy*, vol. 87, no. 1, pp. 197–204, 2010.

[468] S. Jurado, À. Nebot, F. Mugica, and N. Avellana, "Hybrid methodologies for electricity load forecasting: Entropy-based feature selection with machine learning and soft computing techniques," *Energy*, vol. 86, pp. 276–291, 2015.

[469] I. Koprinska, M. Rana, and V. G. Agelidis, "Yearly and seasonal models for electricity load forecasting," in *Proceedings of the International Joint Conference on Neural Networks*, 2011, pp. 1474–1481.

[470] J. L. Mathieu, P. N. Price, S. Kiliccote, and M. A. Piette, "Quantifying changes in building electricity use, with application to demand response," *IEEE Trans. Smart Grid*, vol. 2, no. 3, pp. 507–518, Sep. 2011.

[471] F. McLoughlin, A. Duffy, and M. Conlon, "Evaluation of time series techniques to characterise domestic electricity demand," *Energy*, vol. 50, no. 1, pp. 120–130, 2013.

[472] F. McLoughlin, A. Duffy, and M. Conlon, "Characterising domestic electricity consumption patterns by dwelling and occupant socio-economic variables: An Irish case study," *Energy Build.*, vol. 48, no. July 2009, pp. 240–248, 2012.

[473] R. Mena, F. Rodríguez, M. Castilla, and M. R. Arahal, "A prediction model based on neural networks for the energy consumption of a bioclimatic building," *Energy Build.*, vol. 82, pp. 142–155, 2014.

[474] G. R. Newsham and B. J. Birt, "Building-level occupancy data to improve ARIMA-based electricity use forecasts," in *Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building - BuildSys '10*, 2010, p. 13.

[475] I. Richardson, M. Thomson, D. Infield, and C. Clifford, "Domestic electricity use: A high-resolution energy demand model," *Energy Build.*, vol. 42, no. 10, pp. 1878–1887, 2010.

[476] L. G. Swan, V. I. Ugursal, and I. Beausoleil-Morrison, "Occupant related household energy consumption in Canada: Estimation using a bottom-up neural-network technique," *Energy Build.*, vol. 43, no. 2–3, pp. 326–337, 2011.

[477] J. Widén and E. Wäckelgård, "A high-resolution stochastic model of domestic activity patterns and electricity demand," *Appl. Energy*, vol. 87, no. 6, pp. 1880–1892, 2010.

[478] J. Yoo and K. Hur, "Load forecast model switching scheme for improved robustness to changes in building energy consumption patterns," *Energies*, vol. 6, no. 3, pp. 1329–1343, Mar. 2013.

[479] G. Zahedi, S. Azizi, A. Bahadori, A. Elkamel, and S. R. Wan Alwi, "Electricity demand estimation using an adaptive neuro-fuzzy network: A case study from the Ontario province - Canada," *Energy*, vol. 49, no. 1, pp. 323–328, 2013.

[480] S. F. Crone, J. Guajardo, and R. Weber, "The impact of preprocessing on support vector regression and neural networks in time series prediction," in *Proceedings of the 2006 International Conference on Data Mining, DMIN*, 2006, pp. 37–44.

[481] J. Huang, Y.-F. Li, and M. Xie, "An empirical analysis of data preprocessing for machine learning-based software cost estimation," *Inf. Softw. Technol.*, vol. 67, pp. 108–127, 2015.

[482] Suhartono and Subanar, "The Effect of Decomposition Method As Data Preprocessing on Neural Networks Model for forecasting trend and seasonal time series," *J. Tek. Ind.*, vol. 8, no. 2, pp. 156–164, 2006.

[483] S. Technology, "Semantic Technology and Linked Data Annotation," *Language (Baltim).*

[484] C. J. Anumba, J. Pan, R. R. A. Issa, and I. Mutis, "Collaborative project information management in a semantic web environment," *Eng. Constr. Archit. Manag.*, vol. 15, no. 1, pp. 78–94, Jan. 2008.

[485] M. D'Aquin, J. Davies, and E. Motta, "Smart Cities' Data: Challenges and Opportunities for Semantic Technologies," *IEEE Internet Comput.*, vol. 19, no. 6, pp. 66–70, 2015.

[486] A. Bryman, *Socail research methods*. 2012.

[487] M. Saunders, P. Lewis, and A. Thornhill, "Research Methods for Business Students," *Res. methods Bus. students*, p. 649, 2008.

[488] E. G. Guba and Y. S. Lincoln, "Epistemological and methodological bases of naturalistic inquiry," *Educ. Commun. Technol.*, vol. 30, no. 4, pp. 233–252, 1982.

[489] W. Filstead, "Qualitative methods: A needed perspective in evaluation research," in *Qualitative and quantitative methods in evaluation research*, 1979, pp. 33–48.

[490] Susweta GuhaThakurta, "Understanding research philosophy," *Project Guru*, 2015. [Online]. Available: https://www.projectguru.in/publications/research-philosophy/. [Accessed: 20-Jun-2018].

[491] J. Dudovskiy, *The Ultimate Guide to Writing a Dissertation in Business Studies: A Step-by-Step Assistance*. 2018.

[492] M. Tedre, "Know Your Discipline: Teaching the Philosophy of Computer Science," *J. Inf. Technol. Educ. Res.*, vol. 6, no. 1, pp. 105–122, 2017.

[493] R. Turner and A. H. Eden, "The Philosophy of Computer Science," *J. Appl. Log.*, vol. 6, no. 4, p. 459, 2008.

[494] S. Kemmis, R. McTaggart, and R. Nixon, *The Action Research Planner*, no. Carr 2006. Singapore: Springer Singapore, 2013.

[495] R. L. Baskerville, "Investigating Information Systems with Action Research," *Commun. Assoc. Inf. Syst.*, vol. 2, 2018.

[496] M. D. Myers and D. E. Avison, "Qualitative Research in Information Systems: A Reader," in *Introducing Qualitative Methods Series*, 2002, pp. 3–25.

[497] S. Wang and H. Wang, "Towards Innovative Design Research in Information Systems," *J. Comput. Inf. Syst.*, vol. 51, no. 1, pp. 11–18, 2010.

[498] J. W. Creswell, V. L. P. Clark, and V. L. Plano Clark, *Designing and Conducting Mixed Methods Research*, vol. 31, no. 4. 2007.

[499] S. O. Migiro and B. a Magangi, "Mixed methods : A review of literature and the future of the new research paradigm," *African J. Bus. Manag.*, vol. 5, no. 10, pp.

3757–3764, 2011.

[500] WHO Regional Office for the Western Pacific, *Health Research Methodology: A Guide for Training in Research Methods*. 2001.

[501] D. A. Grimes and K. F. Schulz, “Descriptive studies: What they can and cannot do,” *Lancet*, vol. 359, no. 9301, pp. 145–149, Jan. 2002.

[502] Alan R. Hevner, “A Three Cycle View of Design Science Research,” *Scand. J. Inf. Syst.*, vol. 19, no. 2, 2017.

[503] D. Tranfield, D. Denyer, and P. Smart, “Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review,” *Br. J. Manag.*, vol. 14, no. 3, pp. 207–222, 2003.

[504] A. W. Righi, T. A. Saurin, and P. Wachs, “A systematic literature review of resilience engineering: Research areas and a research agenda proposal,” *Reliab. Eng. Syst. Saf.*, vol. 141, pp. 142–152, 2015.

[505] J. P. T. Higgins and S. Green (editors), *Cochrane Handbook for Systematic Reviews of Interventions*, The Cochra., no. 4. 2006.

[506] D. A. Kolb, “Experiential Learning: Experience as the Source of Learning and Development,” *Prentice Hall, Inc.*, vol. 8, no. 1984, pp. 20–38, 1984.

[507] THERMOSS project consortium, “THERMOSS,” 2018. [Online]. Available: https://thermoss.eu/. [Accessed: 27-Jun-2018].

[508] Pentagon Project Consortium, “PENTAGON - Pentagon Project,” 2018. [Online]. Available: http://www.pentagon-project.eu/. [Accessed: 29-Jun-2018].

[509] MAS^{2}TERING Consortium, “Home - MAS^{2}TERING,” 2018. [Online]. Available: http://www.mas2tering.eu/. [Accessed: 29-Jun-2018].

[510] FIPA-ACL, “FIPA Agent Communication Language Overview.” 2000.

[511] International Electrotechnical Commission, “IEC - Smart grid Standards,” 2018. [Online]. Available: http://www.iec.ch/smartgrid/standards/. [Accessed: 05-Jul-2018].

[512] openADR alliance, “openADR,” 2018. [Online]. Available: https://www.openadr.org/. [Accessed: 05-Jul-2018].

[513] H. A. Akkermans, P. Bogerd, E. Yücesan, and L. N. van Wassenhove, “The impact of ERP on supply chain management: Exploratory findings from a European Delphi study,” *Eur. J. Oper. Res.*, vol. 146, no. 2, pp. 284–301, Apr. 2003.

[514] G. J. Skulmoski and F. T. Hartman, “The Delphi Method for Graduate Research,” *J. Inf. Technol. Educ.*, vol. 6, no. 1, pp. 1–21, 2007.

[515] E. De Vet, J. Brug, J. De Nooijer, A. Dijkstra, and N. De Vries, “Determinants of forward stage transitions: A Delphi study,” *Health Educ. Res.*, vol. 20, no. 2, pp. 195–205, 2005.

[516] R. Loo, “The Delphi method: A powerful tool for strategic management,” *Policing*, vol. 25, no. 4, pp. 762–769, 2002.

[517] S. A. Alshehri, “A Proposed Framework for Resilience to Biological Disasters: The Case of MERS-COV Threat in a Transient Mass Gathering Event,” 2016.

[518] S. J. van Zolingen and C. A. Klaassen, “Selection processes in a Delphi study about key qualifications in Senior Secondary Vocational Education,” *Technol. Forecast. Soc. Change*, vol. 70, no. 4, pp. 317–340, May 2003.

[519] Jisc, “Online surveys,” 2018. [Online]. Available: https://www.onlinesurveys.ac.uk/. [Accessed: 12-Jul-2018].

[520] K. Janowicz, “Observation-Driven Geo-Ontology Engineering,” *Trans. GIS*, vol. 16, no. 3, pp. 351–374, 2012.

[521] M. Hristozova and L. Sterling, “Experiences with Ontology Development for Value-Added Publishing,” *3rd Work. Ontol. Agent Syst.*, p. 17, 2003.

[522] J. Garrido and I. Requena, “Proposal of ontology for environmental impact assessment: An application with knowledge mobilization,” *Expert Syst. Appl.*, vol. 38, no. 3, pp. 2462–2472, 2011.

[523] C. Baccigalupo and E. Plaza, “Poolcasting: A social Web radio architecture for group customisation,” *Proc. - 3rd Int. Conf. Autom. Prod. Cross Media Content Multi-Channel Distrib. AXMEDIS 2007*, vol. SS-97-06, pp. 115–122, 2007.

[524] S. Staab, R. Studer, H. P. Schnurr, and Y. Sure, “Knowledge processes and ontologies,” *IEEE Intell. Syst. Their Appl.*, vol. 16, no. 1, pp. 26–34, 2001.

[525] M. Grüninger, M. S. Fox, and M. Gruninger, “Methodology for the Design and Evaluation of Ontologies,” *Int. Jt. Conf. Artif. Intel. (IJCAI95), Work. Basic Ontol. Issues Knowl. Shar.*, pp. 1–10, 1995.

[526] Swoogle, “Swoogle Semantic Web Search Engine,” 2011. [Online]. Available: http://swoogle.umbc.edu/2006/. [Accessed: 26-Feb-2018].

[527] Watson, “Watson Semantic Web Search,” 2012. [Online]. Available: http://watson.kmi.open.ac.uk/WatsonWUI/. [Accessed: 26-Feb-2018].

[528] “DAML Ontology Library,” 2011. [Online]. Available: http://www.daml.org/ontologies/. [Accessed: 26-Feb-2018].

[529] Stanford, “Protege Ontology Library,” vol. 2010, no. Nov. 2010.

[530] A. Pease, I. Niles, and J. Li, “The Suggested Upper Merged Ontology : A Large Ontology for the Semantic Web and its Applications,” *AAAI Technical Report*, vol. 28, pp. 7–10, 2002.

[531] A. Gangemi, N. Guarino, A. Oltramari, L. Schneider, and C. Masolo, “Sweetening Ontologies with DOLCE,” 2007, pp. 166–181.

[532] R. Arp, B. Smith, and A. D. Spear, *Building Ontologies With Basic Formal Ontology*. 2016.

[533] M. T. Biagetti and M. T. Biagetti, “An ontological model for the integration of cultural heritage information: CIDOC-CRM,” *JLIS.it*, vol. 7, no. 3, pp. 43–77, Sep. 2016.

[534] T. Morris and S. Wood, “Testing the survey method: Continuity and change in british industrial relations,” *Work Employ. Soc.*, vol. 5, no. 2, pp. 259–282, Jun. 1991.

[535] KairosDB, “KairosDB,” 2018. [Online]. Available: http://kairosdb.github.io/. [Accessed: 21-Dec-2018].

[536] M. C. Loui and K. W. Miller, “Ethics and Professional Responsibility in Computing,” *Wiley Encycl. Comput. Sci. Eng.*, pp. 1–12, 2008.

[537] M. A. Pierce and J. W. Henry, “Computer ethics: The role of personal, informal, and formal codes,” *J. Bus. Ethics*, vol. 15, no. 4, pp. 425–437, Apr. 1996.

[538] Cardiff University, “Research Integrity & Governance Code of Practice,” 2017.

[539] M. S. Chandurkar, S. Mugade, S. Sinha, M. Misal, and P. Borekar, “Implementation of Real Time Bus Monitoring and Passenger Information System,” *Int. J. Sci. Res. Publ.*, vol. 3, no. 1, pp. 2250–3153, 2013.

[540] L. Khoudour, “Real-time passenger counting in buses using dense stereovision,” *J. Electron. Imaging*, vol. 19, no. 3, p. 031202, 2010.

[541] M. P. Pelletier, M. Trépanier, and C. Morency, “Smart card data use in public transit: A literature review,” *Transp. Res. Part C Emerg. Technol.*, vol. 19, no. 4, pp. 557–568, Aug. 2011.

[542] H. Ki, K. Soo, B. Technology, and S. Korea, “Development of Road Traffic Analysis Platform Using Big Data,” pp. 65–66.

[543] National Center for Sustainable Transportation, “Quantifying Vehicle Miles Traveled.” National Center for Sustainable Transportation, 2017.

[544] C. Ahn and H. Peng, “Decentralized voltage control to minimize distribution power loss of microgrids,” *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1297–1304, Sep. 2013.

[545] V. Masatin, E. Latõšev, and A. Volkova, “Evaluation Factor for District Heating Network Heat Loss with Respect to Network Geometry,” in *Energy Procedia*, 2016, vol. 95, pp. 279–285.

[546] N. Wain and B. Ariel, “Tracking of police patrol,” *Polic.*, vol. 8, no. 3, pp. 274–283, Sep. 2014.

[547] T. P. Lambrou, C. C. Anastasiou, C. G. Panayiotou, and M. M. Polycarpou, “A low-cost sensor network for real-time monitoring and contamination detection in drinking water distribution systems,” *IEEE Sens. J.*, vol. 14, no. 8, pp. 2765–2772, 2014.

[548] T. J. Mumanga, R. M. Martinez, and M. F. Grobler, “Smart emergency response interface for fire detection,” in *IEEE AFRICON Conference*, 2015, vol. 2015–Novem, pp. 1–4.

[549] American Society of Heating Refrigerating and Air-Conditioning Engineers Inc., “ANSI/ASHRAE Standard 55.” 2010.

[550] Autodesk, “Human Thermal Comfort | Sustainability Workshop.” [Online]. Available: https://sustainabilityworkshop.autodesk.com/buildings/human-thermal-comfort. [Accessed: 15-May-2018].

[551] SmartBin, “Smartbin |Smart City|,” 2017. [Online]. Available: https://www.smartbin.com/. [Accessed: 20-Sep-2017].

[552] H. Akbari, R. Levinson, and P. Berdahl, “ASTM Standards for Measuring Solar Reflectance and Infrared Emittance of Construction Materials and Comparing their Steady-State Surface Temperatures,” *ACEEE Summer Study Energy Effic. Build.*, vol. 1, no. 1, pp. 1–9, 1996.

[553] A. Borisuit, J.-L. Scartezzini, L. Deschamps, M. Münch, and J. Kämpf, “A new device for dynamic luminance mapping and glare risk assessment in buildings,” *Nonimaging Opt. Effic. Des. Illum. Sol. Conc. IX*, vol. 8485, no. September 2017, p. 84850M, 2012.

[554] U. Dalhousie, “Mobility and Governement: Opportunities and Challenges for Service Delivery and Information Management,” 2014.

[555] C. Walker, C. Bueter, A. Hurst, V. Patten, M. Berendsen, and K. Patten, “GLOBE at Night: Using Sky Quality Meters to Measure Sky Brightness.” 2011.

[556] P. Kumar *et al.*, “Real-time sensors for indoor air monitoring and challenges ahead in deploying them to urban buildings,” *Science of the Total Environment*, vol. 560–561. Elsevier, pp. 150–159, 01-Aug-2016.

[557] S. L. Pimm *et al.*, “Emerging Technologies to Conserve Biodiversity,” *Trends in Ecology and Evolution*, vol. 30, no. 11. Elsevier Current Trends, pp. 685–696, 01-Nov-2015.

[558] G. S. Thakur *et al.*, “Real-time Urban Population Monitoring Using Pervasive Sensor Network.”

[559] L. K. Wiginton, H. T. Nguyen, and J. M. Pearce, “Quantifying rooftop solar photovoltaic potential for regional renewable energy policy,” *Comput. Environ. Urban Syst.*, vol. 34, no. 4, pp. 345–357, 2010.

[560] Q. Weng, “Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends,” *Remote Sens. Environ.*, vol. 117, pp. 34–49, 2012.

[561] R. Behling, M. Bochow, S. Foerster, S. Roessner, and H. Kaufmann, “Automated GIS-based derivation of urban ecological indicators using hyperspectral remote sensing and height information,” *Ecol. Indic.*, vol. 48, pp. 218–234, 2015.

[562] W. G. M. Bastiaanssen, D. J. Molden, and I. W. Makin, “Remote sensing for irrigated agriculture: Examples from research and possible applications,” *Agric. Water Manag.*, vol. 46, no. 2, pp. 137–155, Dec. 2000.

[563] H. Li, L. Zheng, Y. Lei, C. Li, Z. Liu, and S. Zhang, “Estimation of water consumption and crop water productivity of winter wheat in North China Plain using remote sensing technology,” *Agric. Water Manag.*, vol. 95, no. 11, pp. 1271–1278, Nov. 2008.

[564] A. Mancini, E. Frontoni, and P. Zingaretti, “Automatic road object extraction from Mobile Mapping Systems,” *Proc. 2012 8th IEEE/ASME Int. Conf. Mechatron. Embed. Syst. Appl. MESA 2012*, pp. 281–286, 2012.

[565] ArcGIS, “Street Furniture Inventory | ArcGIS for Local Government.” [Online]. Available: http://solutions.arcgis.com/local-government/help/street-furniture-

inventory/. [Accessed: 12-Sep-2017].

[566] cityGML, “What is CityGML?” [Online]. Available: https://www.citygml.org/about/. [Accessed: 12-Sep-2017].

[567] NIST, “Toward Real-Time Measurements for Real-Time Decision Making | NIST,” 2016. [Online]. Available: https://www.nist.gov/news-events/news/2016/07/toward-real-time-measurements-real-time-decision-making. [Accessed: 28-Sep-2017].

[568] J. Gong, J. Geng, and Z. Chen, “Real-time GIS data model and sensor web service platform for environmental data management,” *Int. J. Health Geogr.*, vol. 14, no. 1, p. 2, 2015.

[569] M. . Kavyashree and H. Ramesh, “Wetland Mapping and Change Detection Using Remote Sensing and GIS,” *Ijesc*, vol. 6, no. 8, pp. 2356–2359, 2016.

[570] H. M. Pham, Y. Yamaguchi, and T. Q. Bui, “A case study on the relation between city planning and urban growth using remote sensing and spatial metrics,” *Landsc. Urban Plan.*, vol. 100, no. 3, pp. 223–230, 2011.

[571] C. Wirion, “Using remote sensing to describe urban surface properties for improved hydrological modelling,” no. September, pp. 1–9, 2015.

[572] I. Curebal, R. Efe, H. Ozdemir, A. Soykan, and S. Sönmez, “GIS-based approach for flood analysis: case study of Keçidere flash flood event (Turkey),” *Geocarto Int.*, vol. 31, no. 4, pp. 355–366, 2016.

[573] R. K. Jaiswal, S. Mukherjee, K. D. Raju, and R. Saxena, “Forest fire risk zone mapping from satellite imagery and GIS,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 4, no. 1, pp. 1–10, Aug. 2002.

[574] S. Lee, J. Choi, and K. Min, “Probabilistic landslide hazard mapping using GIS and remote sensing data at Boun, Korea,” *Int. J. Remote Sens.*, vol. 25, no. 11, pp. 2037–2052, 2004.

[575] U. Bhangale, V. Rathod, N. Rajgor, J. Rami, and K. Kurte, “Identification of informal settlement using Remote Sensing Images,” in *Proceedings of the International Conference on Advances in Information Communication Technology & Computing - AICTC ’16*, 2016, pp. 1–5.

[576] World bank group, “Global Tracking Framework 2017: Progress Towards Sustainable Energy.” Washington, DC, 2017.

[577] T. Vivian W.Y., “Rate of Reusable and Recyclable Waste in Construction,” *Open Waste Manag. J.*, vol. 4, no. 1, pp. 28–32, 2011.

[578] B. Davis, P. Baxandall, and R. Pierannunzi, “Transparency in City Spending,” 2013.

[579] W. Raghupathi and V. Raghupathi, “Big data analytics in healthcare: promise and potential,” *Heal. Inf. Sci. Syst.*, vol. 2, no. 1, 2014.

[580] Autodesk Knowledge Network, “Material quantities | Revit Products,” 2017. [Online]. Available: https://knowledge.autodesk.com/support/revit-products/learn-explore/caas/CloudHelp/cloudhelp/2014/ENU/Revit/files/GUID-101E6FBF-D776-

4CFA-A40D-D2DB7C0BDE9A-htm.html. [Accessed: 21-Sep-2017].

[581] S. Wang, M. J. Paul, and M. Dredze, “Social media as a sensor of air quality and public response in China,” *J. Med. Internet Res.*, vol. 17, no. 3, p. e22, Mar. 2015.

[582] D. Agostino and M. Arnaboldi, “Social media data used in the measurement of public services effectiveness: Empirical evidence from Twitter in higher education institutions,” *Public Policy Adm.*, vol. 32, no. 4, pp. 296–322, Oct. 2017.

[583] M. Argiolas, “Surveying housing market supply affordability using a spatial data mining approach,” in *WIT Transactions on Ecology and the Environment*, 2014, vol. 191, pp. 125–137.

[584] V. Landeiro, D. Reis, and A. Culotta, “Using Matched Samples to Estimate the Effects of Exercise on Mental Health from Twitter,” in *AAAI*, 2015.

[585] Y. Kang, J. Park, and A. Kang, “An Analysis on the Spatial Characteristics of Satisfaction on the Residential Environment Using Tweets,” *Int. J. Geospatial Environ. Res.*, vol. 1, no. 2, 2015.

[586] J. Humphry, *Homeless and Connected: Mobile phones and the Internet in the lives of homeless Australians*. 2014.

[587] T. Walk-Morri, “Can Mobile Apps Help the Homeless?,” *Pacific Standard*, 2017. [Online]. Available: https://psmag.com/news/can-mobile-apps-help-the-homeless. [Accessed: 15-May-2018].

[588] Keeney Sinead, Hasson Felicity, and McKenna Hugh, “Consulting the oracle: ten lessons from using the Delphi technique in nursing research,” *J. Adv. Nurs.* , vol. 52, no. 2, pp. 205–12, 2006.

[589] D. Gryth, L. Svensson, A. Rüter, M. Jirwe, M. Castrén, and M. Rådestad, “Essential key indicators for disaster medical response suggested to be included in a national uniform protocol for documentation of major incidents: a Delphi study,” *Scand. J. Trauma. Resusc. Emerg. Med.*, vol. 21, no. 1, pp. 1–11, 2013.

[590] M. Jirwe, K. Gerrish, S. Keeney, and A. Emami, “Identifying the core components of cultural competence: Findings from a Delphi study,” *J. Clin. Nurs.*, vol. 18, no. 18, pp. 2622–2634, 2009.

[591] M. J. Clayton, “Delphi: a technique to harness expert opinion for critical decision-making tasks in education,” *Educ. Psychol.*, vol. 17, no. 4, pp. 373–386, 2007.

[592] H. P. McKenna, “The Delphi technique: a worthwhile research approach for nursing?,” *J. Adv. Nurs.*, vol. 19, no. 6, pp. 1221–1225, 1994.

[593] E. de Leeuw and W. de Heer, “Trends in Household Survey Nonresponse: A Longitudinal and International Comparison,” *Surv. nonresponse*, no. December 2015, pp. 41–54, 2002.

[594] K. M. Hansen, “The effects of incentives, interview length, and interviewer characteristics on response rates in a CATI-study,” *Int. J. Public Opin. Res.*, vol. 19, no. 1, pp. 112–121, Mar. 2007.

[595] R. Curtin, S. Presser, and E. Singer, “Changes in telephone survey nonresponse over

the past quarter century," *Public Opin. Q.*, vol. 69, no. 1, pp. 87–98, 2005.

[596] J. Petchenik and D. J. Watermolen, "A Cautionary Note on Using the Internet to Survey Recent Hunter Education Graduates," *Hum. Dimens. Wildl.*, vol. 16, no. 3, pp. 216–218, May 2011.

[597] K. L. Manfreda, J. Berzelak, V. Vehovar, M. Bosnjak, and I. Haas, "Web Surveys versus other Survey Modes: A Meta-Analysis Comparing Response Rates," *Int. J. Mark. Res.*, vol. 50, no. 1, pp. 79–104, Jan. 2008.

[598] E. Singer and M. P. Couper, "Do Incentives Exert Undue Influence on Survey Participation? Experimental Evidence," *J. Empir. Res. Hum. Res. Ethics*, vol. 3, no. 3, pp. 49–56, 2008.

[599] E. Singer, R. M. Groves, and A. D. Corning, "American Association for Public Opinion Research Differential Incentives: Beliefs About Practices, Perceptions of Equity, and Effects on Survey Participation," *Source Public Opin. Q.*, vol. 63, no. 2, pp. 251–260, 1999.

[600] M. J. Pedersen and C. V. Nielsen, "Improving Survey Response Rates in Online Panels," *Soc. Sci. Comput. Rev.*, vol. 34, no. 2, pp. 229–243, Apr. 2016.

[601] T. A. Heberlein and R. Baumgartner, "Factors Affecting Response Rates to Mailed Questionnaires: A Quantitative Analysis of the Published Literature," *Am. Sociol. Rev.*, vol. 43, no. 4, p. 447, 1978.

[602] D. Andrews, B. Nonnecke, and J. Preece, "Electronic Survey Methodology: A Case Study in Reaching Hard-to-Involve Internet Users," *Int. J. Hum. Comput. Interact.*, vol. 16, no. 2, pp. 185–210, Oct. 2003.

[603] S. Dahlberg, "Web-based expert surveys," *QoG Work. Pap. Ser.*, no. November, 2007.

[604] Welsh Governement, "Local authority municipal waste management," 2018. [Online]. Available: http://gov.wales/statistics-and-research/local-authority-municipal-waste-management/?lang=en.

[605] UK Government-Department for Business Energy & Industrial Strategy, "Conversion factors 2017: Condensed set for most users v02-00." 2017.

[606] DECC and CHP Focus, "CHP – Environmental," pp. 1–24, 2008.

[607] Energy Saving Trust, "Our Calculations," 2016. [Online]. Available: http://www.energysavingtrust.org.uk/about-us/our-calculations. [Accessed: 12-Nov-2018].

[608] M. Presser and M. Larranaga, "URB-grade decision support tool," in *2015 International Conference on Recent Advances in Internet of Things, RIoT 2015*, 2015, no. April.

[609] N. Zhang, H. Chen, X. Chen, and J. Chen, "Semantic framework of internet of things for smart cities: Case studies," *Sensors (Switzerland)*, vol. 16, no. 9, 2016.

[610] M. Kettouch, C. Luca, O. Khorief, R. Wu, and S. Dascalu, "Semantic data management in smart cities," *Proc. - 2017 Int. Conf. Optim. Electr. Electron.*

Equipment, OPTIM 2017 2017 Intl Aegean Conf. Electr. Mach. Power Electron. ACEMP 2017, pp. 1126–1131, 2017.

[611] S. Bischof, A. Karapantelakis, A. Sheth, and A. Mileo, “Semantic Modelling of Smart City Data Description of Smart City Data,” *W3C Work. Web Things Enablers Serv. an open Web Devices*, pp. 1–5, 2014.

[612] C. M. Chituc, C. Toscano, and A. L. Azevedo, “Towards seamless interoperability in collaborative networks,” *IFIP Int. Fed. Inf. Process.*, vol. 243, pp. 445–452, 2007.

[613] I. Alvertis, D. Askounis, M. Petychakis, R. Tsouroplis, F. Lampathaki, and E. Biliri, “Enterprise Collaboration Framework for Managing, Advancing and Unifying the Functionality of Multiple Cloud-Based Services with the Help of a Graph API,” 2014, pp. 153–160.

[614] M. Dibley, H. Li, Y. Rezgui, and J. Miles, “An ontology framework for intelligent sensor-based building monitoring,” *Autom. Constr.*, vol. 28, pp. 1–14, Dec. 2012.

[615] S. J. D. Cox, “Ontology for observations and sampling features, with alignments to existing models,” *Semant. Web*, vol. 8, no. 3, pp. 453–470, 2017.

[616] S. Cox, “Observations and Measuremets-XML Implementation,” *Measurement*, pp. 1–66, 2011.

[617] T. Liebich *et al.*, “Industry Foundation Classes Release 4 (IFC4),” 2013. [Online]. Available: http://www.buildingsmart-tech.org/ifc/IFC4/final/html/.

[618] P. Pauwels and A. Roxin, “SimpleBIM : From full ifcOWL graphs to simplified building graphs,” *11th Eur. Conf. Prod. Process Model.*, no. October, pp. 11–18, 2016.

[619] D. Brickley and L. Miller, “The FOAF Project,” 2015. [Online]. Available: http://www.foaf-project.org/. [Accessed: 26-Nov-2018].

[620] D. Calvanese *et al.*, “Optique: OBDA solution for big data,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7955 LNCS, pp. 293–295, 2013.

[621] N. Konstantinou and D. E. Spanos, *Materializing the web of linked data*. 2015.

[622] E. Kharlamov *et al.*, “Ontology-Based Integration of Streaming and Static Relational Data with Optique,” in *Proceedings of the 2016 International Conference on Management of Data - SIGMOD ’16*, 2016, no. ii, pp. 2109–2112.

[623] D. Calvanese *et al.*, “The optique project: Towards OBDA systems for industry,” *CEUR Workshop Proc.*, vol. 1080, pp. 7–11, 2013.

[624] M. Rodriguez-Muro, R. Kontchakov, and M. Zakharyaschev, “OBDA with Ontop,” in *CEUR Workshop Proceedings*, 2013, vol. 1015, pp. 128–133.

[625] H. J. Ter Horst, “Completeness, decidability and complexity of entailment for RDF Schema and a semantic extension involving the OWL vocabulary,” *Web Semant.*, vol. 3, no. 2–3, pp. 79–115, 2005.

[626] J. Pokorný, “Database technologies in the world of big data,” 2015, pp. 1–12.

[627] T. M. Cannolly and C. E. Begg, *Database Systems : A Practical Approach to design,*

Implementation, and Managemnet(5th ed.). 2005.

[628] S. S. Harkins and M. W. P. Reid, “An Introduction to Relational Database Theory,” in *SQL: Access to SQL Server*, 2013, pp. 35–68.

[629] A. Salehnia, “Comparisions of relational Databases with Big Data: a Teaching Approach,” *ACM Comput. Surv.*, vol. 8, no. I, pp. 43–66, 2015.

[630] Oracle Corporation, “MySQL,” 2019. [Online]. Available: https://www.mysql.com/fr/. [Accessed: 18-Feb-2019].

[631] “PostgreSQL: The world’s most advanced open source database,” 2016. [Online]. Available: https://www.postgresql.org/. [Accessed: 18-Feb-2019].

[632] Microsoft, “Microsoft Data Platform,” 2019. [Online]. Available: https://www.microsoft.com/en-us/sql-server/. [Accessed: 18-Feb-2019].

[633] IBM, “IBM Db2 – Data management software – IBM Analytics,” 2019. [Online]. Available: https://www.ibm.com/analytics/us/en/db2/. [Accessed: 18-Feb-2019].

[634] 52°North, “Home - 52°North Initiative for Geospatial Open Source Software GmbH,” 2015. [Online]. Available: http://52north.org/. [Accessed: 20-Dec-2018].

[635] A. Bröring, C. Stasch, and J. Echterhoff, “OGC Sensor Observation Service,” *OGC Implement. Stand.*, vol. 12-006, p. 163, 2012.

[636] M. Botts, G. Percivall, C. Reed, and J. Davidson, “OGC® sensor web enablement: Overview and high level architecture,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 4540 LNCS, no. December, pp. 175–190, 2008.

[637] Open Geospatial Consortium, “OGC ® OWS-6 SensorML Profile for Discovery Engineering Report.” 2009.

[638] G. Percivall, “Open Geospatial Consortium OGC Smart Cities Spatial Information Framework,” pp. 1–32, 2015.

[639] “52°North Sensor Observation Service 4.x Database model,” 2018. [Online]. Available: https://wiki.52north.org/SensorWeb/SensorObservationServiceDatabaseModel. [Accessed: 20-Dec-2018].

[640] Hibernate, “Hibernate. Everything data.” [Online]. Available: http://hibernate.org/. [Accessed: 21-Dec-2018].

[641] A. C. Rai *et al.*, “End-user perspective of low-cost sensors for outdoor air pollution monitoring,” *Science of the Total Environment*, vol. 607–608. Elsevier, pp. 691–705, 31-Dec-2017.

[642] Stanford, “Protégé,” 2019. [Online]. Available: https://protege.stanford.edu/. [Accessed: 21-Dec-2018].

[643] Pieter Pauwels, “How can I efficiently use ifcOWL-based RDF graphs? – openBIMstandards.org.” .

[644] RML Editor, “RML Generic Mapping Language.” [Online]. Available: http://rml.io/.

[Accessed: 22-Dec-2018].

[645] P. Heyvaert, A. Dimou, R. Verborgh, and E. Mannens, “Semi-automatic example-driven Linked Data mapping creation,” in *CEUR Workshop Proceedings*, 2017, vol. 1946, pp. 14–25.

[646] A. Dimou, M. Vander Sande, P. Colpaert, R. Verborgh, E. Mannens, and R. Van De Walle, “RML: A generic language for integrated RDF mappings of heterogeneous data,” in *CEUR Workshop Proceedings*, 2014, vol. 1184.

[647] G. M. Santipantakis, K. I. Kotis, G. A. Vouros, and C. Doulkeridis, “RDF-Gen,” in *Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics - WIMS ’18*, 2018, pp. 1–10.

[648] D. Ulutaş Karakol, G. Kara, C. Yılmaz, and C. Cömert, “Semantic linking spatial rdf data to the web data sources,” in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 2018, vol. 42, no. 4, pp. 711–715.

[649] GeoTriples, “GeoTriples.” [Online]. Available: http://geotriples.di.uoa.gr/. [Accessed: 22-Dec-2018].

[650] D. Calvanese, B. Cogrel, S. Komla-ebri, R. Kontchakov, and D. Lanti, “Ontop : Answering SPARQL Queries over Relational Databases,” 2017.

[651] K. Bereta, G. Xiao, and M. Koubarakis, “Answering geosparql queries over relational data,” *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.*, vol. 42, no. 4W2, pp. 43–50, 2017.

[652] K. Bereta and M. Koubarakis, “Ontop of geospatial databases,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9981 LNCS, no. 4, 2016, pp. 37–52.

[653] K. Bereta, G. Xiao, M. Koubarakis, M. Hodrius, C. Bielski, and G. Zeug, “Ontop-spatial: Geospatial data integration using GeoSPARQL-to-SQL translation,” in *CEUR Workshop Proceedings*, 2016, vol. 1690, no. 4, pp. 331–342.

[654] M. Kintz, “A Semantic Dashboard Description Language for a Process-oriented Dashboard Design Methodology,” *CEUR Workshop Proc.*, vol. 947, pp. 31–36, 2012.

[655] R. Carli, V. Albino, M. Dotoli, G. Mummolo, and M. Savino, “A dashboard and decision support tool for the energy governance of smart cities,” *2015 IEEE Work. Environ. Energy, Struct. Monit. Syst. EESMS 2015 - Proc.*, pp. 23–28, 2015.

[656] V. Kumaraswamy, S. Ergan, and B. Akinci, “A Taxonomy for Building Energy Dashboards,” 2014, pp. 2192–2199.

[657] A. Cahyadi and A. Prananto, “Reflecting design thinking: A case study of the process of designing dashboards,” *J. Syst. Inf. Technol.*, vol. 17, no. 3, pp. 296–306, 2015.

[658] R. Yun *et al.*, “Toward the design of a dashboard to promote environmentally sustainable behavior among office workers,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7822 LNCS, pp. 246–252, 2013.

[659] K. S. Vallerio, L. Zhong, and N. K. Jha, “Energy-efficient graphical user interface design,” *IEEE Trans. Mob. Comput.*, vol. 5, no. 7, pp. 846–859, 2006.

[660] T. Dkaki and J. Mothe, “An energy-based model to optimize cluster visualization,” *Proc. - Int. Conf. Res. Challenges Inf. Sci.*, 2014.

[661] wikidata.org, “Wikidata Query Service.” [Online]. Available: https://query.wikidata.org/. [Accessed: 28-Jan-2019].

[662] wikidata.org, “Wikidata: Result Views.” [Online]. Available: https://www.wikidata.org/wiki/Wikidata:SPARQL_query_service/Wikidata_Query_Help/Result_Views. [Accessed: 28-Jan-2019].

[663] wikidata.org, “Wikidata: Query Helper.” [Online]. Available: https://www.wikidata.org/wiki/Wikidata:SPARQL_query_service/Query_Helper. [Accessed: 28-Jan-2019].

[664] A. GhaffarianHoseini *et al.*, “Application of nD BIM Integrated Knowledge-based Building Management System (BIM-IKBMS) for inspecting post-construction energy efficiency,” *Renew. Sustain. Energy Rev.*, vol. 72, pp. 935–949, May 2017.

[665] T. Winchcomb, S. Massey, P. Beastall, and OFCOM, “Review of latest developments in the Internet of Things,” 2017.

[666] J. Rice and N. Martin, “Smart infrastructure technologies: Crowdsourcing future development and benefits for Australian communities,” *Technological Forecasting and Social Change*, North-Holland, 10-Apr-2018.

[667] Y. Sure, S. Staab, and R. Studer, “On-To-Knowledge Methodology (OTKM),” in *Handbook on Ontologies*, Springer, 2013, pp. 117–132.

[668] A. De Nicola, M. Missikoff, and R. Navigli, “A software engineering approach to ontology building,” *Inf. Syst.*, vol. 34, no. 2, pp. 258–275, 2009.

[669] C. Kuster, J. L. Hippolyte, and Y. Rezgui, “Collaborative Network for District Energy Operation and Semantic Technologies: A Case Study,” in *IFIP Advances in Information and Communication Technology*, 2018, vol. 534, pp. 486–495.

[670] The European Parliament and the Council of the European Union, *Regulation (EU) 2016/679 (GDPR)*, vol. 59. 2016, pp. 1–88.

[671] E. Graham-Harrison and C. Cadwalladr, “Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach,” *the Guardian*, 2018. [Online]. Available: https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election. [Accessed: 12-Feb-2019].

[672] A. H. Ewen MacAskill, “Edward Snowden: ‘The people are still powerless, but now they’re aware,’” *The Guardian*, 2018.

[673] D. Arnott and G. Pervan, “A critical analysis of decision support systems research revisited: The rise of design science,” *Enacting Res. Methods Inf. Syst. Vol. 3*, vol. 29, no. 4, pp. 43–103, 2016.

[674] J. C. Hershauer, “The New Science of Management Decision.,” *Academy of Management Review*, vol. 3, no. 1. pp. 161–162, 2011.

[675] S. Liu, A. H. B. Duffy, R. I. Whitfield, and I. M. Boyle, “Integration of decision support systems to improve decision support performance,” *Knowl. Inf. Syst.*, vol. 22, no. 3, pp. 261–286, 2010.

[676] A. Larsson, O. Ibrahim, A. Talantsev, and A. Larsson, “Decision Support Framework,” 2014.

[677] E. Bendoly, *Excel basics to blackbelt: an accelerated guide to decision support designs*, vol. 46, no. 07. 2013.

[678] E. Oliver, A. Taylor, and M. Rocchetti, “MAS2TERING: Internal Note Blaenau Gwent Test Site Description.” 2017.